ON THE MAXIMUM CHI-SQUARED TEST
FOR WHITE NOISE

BY

TERENCE L. SINCICH

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1980

TO MY FAMILY

AND FAITH

ACKNOWLEDGEMENTS

I am indebted to Dr. James T. McClave,not only for his patience and encouragement while guiding me through this dissertation, but also for the invaluable friendship he provided during my course of graduate study. For their helpful suggestions towards this research, I wish to thank Dr. John G. Saw, Dr. Mark C. Yang, Dr. Andre Khuri, Dr. Ramon C. Littel and Dr. Andrew Rosalsky. Also, I am most appreciative of Dr. Dennis D. Wackerly and Dr. Richard L. Scheaffer for their understanding and assistance in matters related to my graduate program.

Finally, I especially thank my typist, Cecily Noble, for her amazing ability to transform the raw text, bulky equations and lengthy tables handed her into immaculate typed copy.

O M45 80. 285.1

TABLE OF CONTENTS

Page
ACKNOWLEDGEMENTS iii

ABSTRACT vi

CHAPTERS

I INTRODUCTION 1

1.1 The Problem 1
1.2 Various Tests 4
1.2.1 The Durbin-Watson Bounds Test 4
1.2.2 Beta Approximation 9
1.2.3 d_u Approximation 10
1.2.4 BLUS Estimators 11
1.2.5 The Sign Test 13
1.2.6 Periodogram Based Tests 14
1.3 Concluding Note 16

II COMPARING TESTS OF SERIAL CORRELATION 17

2.1 Power Comparisons 17
2.1.1 Exact d Test Versus Approximate Tests of Serial Correlation 17
2.1.2 d Test Versus BLUS Test 18
2.1.3 The Durbin-Watson d Test Versus Non-first Order Alternative Tests 19
2.2 The Alternative Error Model 20

III TWO TESTS DESIGNED TO DETECT A GENERAL ALTERNATIVE 22

3.1 Box and Pierce 'Portmanteau' Test 22
3.2 A New Technique: The Max-χ^2 Procedure 25
3.3 Asymptotic Distribution of the Test Statistics 30
3.3.1 Preliminaries 30
3.3.2 General Case 31
3.3.3 Special Cases 42
3.3.3.1 Model 1: the Case J=2, K=2 49
3.3.3.2 Model 2: the Case J=2, K=4 53
3.3.3.2.1 Approximate Asymptotic Power of the Box-Pierce Test. 55
3.3.3.2.2 Approximate Asymptotic Power of the Max-χ^2 Test 60

IV A POWER COMPARISION OF THE BOX-PIERCE 'PORTMANTEAU', MAX-χ^2, AND D TESTS 71

4.1 Monte Carlo Simulations 71

4.1.1 Observable Residuals: No Regression................. 74
4.1.2 Estimated Residuals After Regression................ 95
4.2 Power Approximations for Large n...........................113
4.2.1 Case: J=2, K=4...114
4.2.2 Case: J=3, K=4...115
4.2.3 Case: J=4, K=4...116
4.2.4 Approximation Results..................................117
4.2.5 A Note on the Taylor-Series Approximation...........128
4.3 Summary...131

V OTHER PROPERTIES OF THE BOX-PIERCE AND MAX-χ^2 TESTS.............133

5.1 Asymptotic Powers for Large n..............................133
5.2 Hodges and Lehmann Asymptotic Relative Efficiency..........142
5.3 Likelihood Ratio...153
5.3.1 General Case...154
5.3.2 Special Case: J=5, K=2.................................156
5.3.3 Special Case: J=K......................................158

VI CONCLUSION...164

6.1 Concluding Remarks...164
6.2 Future Research..165

APPENDIX..169

BIBLIOGRAPHY..213

BIOGRAPHICAL SKETCH...216

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

ON THE MAXIMUM CHI-SQUARED TEST FOR WHITE NOISE

By

Terence L. Sincich
August, 1980

Chairman: James T. McClave
Major Department: Statistics

Consider the general regression model

$$Y_t = \beta_o + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_g X_{gt} + Z_t , \quad t = 1, 2, \ldots, n$$

where the dependent variable Y_t and the independent variables $X_{1t}, X_{2t}, \ldots, X_{gt}$ are observed and the vector of model parameters $\underset{\sim}{\beta}' = (\beta_o, \beta_1, \ldots, \beta_g)$ is unknown. The usual assumption on the unobservable errors $\{Z_t\}$ is that $Z_1, Z_2, Z_3, \ldots, Z_n$ are independent with zero mean and constant variance. When this assumption is violated, the application of the least squares regression method will often lead to erroneous inferences on the elements of $\underset{\sim}{\beta}$. In order to avoid this problem, a test for independence of the errors is essential.

Many testing procedures have been proposed for this problem and a summary of the most widely used tests is given. However, these tests attain high power only when the regression errors follow the first order autoregressive process

$$Z_t = \phi Z_{t-1} + \varepsilon_t ,$$

where $|\emptyset|<1$ and $\{\varepsilon_t\}$ is white noise. We consider two tests, based on the vector of sample autocorrelations, which are designed to have high power against a more general autoregressive alternative: the Box-Pierce test and a new procedure called the max-χ^2 test. An extensive Monte-Carlo simulation study is undertaken in order to compare the powers of the tests under the first order lag J autoregressive alternative

$$Z_t = \emptyset Z_{t-J} + \varepsilon_t \quad ,$$

where J is not necessarily 1. This particular model is chosen for it is one which is near white noise, and hence one for which many testing procedures will fail to detect the presence of serially correlated errors. A graphical display of the simulated powers of the tests shows that although both tests attain a high power in this instance, the max-χ^2 test clearly outperforms the Box-Pierce procedure.

The two tests are also compared to the Durbin-Watson d test for several least squares regression models. Since the d test is optimal when J=1 in the alternative error model, it is recommended as the initial test to apply in practice. However, the d test performs poorly for J>1 and should be supplemented by another test which has high power in this case. Power simulation results indicate that the max-X^2 test would be the better supplementary test.

Since the problem of analytically deriving the exact powers of the tests is intractable even for simple alternative error models, asymptotic powers are considered. Approximations to the asymptotic powers of the Box-Pierce and max-X^2 tests are discussed and their performance evaluated.

Comparisons of the Box-Pierce and max-X^2 tests are also made with respect to Hodges-Lehmann efficiency and deficiency and with respect to likelihood ratio. Although the asymptotic relative efficiency of the

tests is 1, the Box-Pierce test is shown to be asymptotically deficient when compared to the max-χ^2 test. A derivation of the likelihood ratio statistic for special cases of the problem reveals that the max-χ^2 test is asymptotically equivalent to the likelihood ratio test.

CHAPTER I

INTRODUCTION

1.1 The Problem

Consider the general regression model

$$Y_t = \beta_o + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_g X_{gt} + Z_t$$

where the dependent variable, Y_t, and the independent variables, X_{1t}, . . ., X_{gt}, are observed and the random error, Z_t, is unobserved, $t = 1, 2, \ldots, n$. In most practical cases, the vector of model parameters, $\underset{\sim}{\beta}' = (\beta_o, \beta_1, \ldots, \beta_g)$ will be unknown. The researcher who is interested in estimating the parameters for prediction purposes usually applies the least squares regression method. The least squares model can be written as follows:

$$\hat{Y}_t = \hat{\beta}_o + \hat{\beta}_1 X_{1t} + \hat{\beta}_2 X_{2t} + \ldots + \hat{\beta}_g X_{gt} \tag{1.1.1}$$

where $\hat{\underset{\sim}{\beta}}' = (\hat{\beta}_o, \hat{\beta}_1, \ldots, \hat{\beta}_g)$ is the vector of parameter estimates and $\hat{Y}_t$ is the predicted value of the dependent variable for the particular set of independent variables $(X_{1t}, X_{2t}, \ldots, X_{gt})$ observed.

Before utilizing the least squares prediction model, the researcher often wishes to make inferences about the components of $\underset{\sim}{\beta}$. To ensure the validity of these inferences, the following assumptions must be made:

(1) The error terms (Z_t's) have zero mean and equal variance, i.e., $Z_t \sim (0,\sigma^2)$.

(2) If the independent variables in the model are random, then the

error terms are distributed independently of these independent variables, i.e., Z_t is distributed independently of the X_t's.

(3) Successive errors are uncorrelated.

When autoregressive schemes or stochastic difference equations are used in the regression model, lagged values of the dependent variable occur as independent variables. These types of models violate assumption (2), and will be excluded from further consideration in this research. For models which do not contain lagged values of the dependent variables as regressors, the independent variables can be regarded as fixed, known constants, even if they are in fact random variables. All inferences are then conditional on the observed values of the independent variables. Of importance, then, is determining whether the model under consideration is in violation of assumption (3). When this assumption is violated, as is often the case in the analysis of time series data and in econometric modeling, complications arise. In order to grasp the severity of the problem, let us consider the following:

Rewrite the regression model (1.1.1) in matrix form as

$$\underset{\sim}{Y}_{(n\times 1)} = X_{(n\times g)}\, \underset{\sim}{\beta}_{(g\times 1)} + \underset{\sim}{Z}_{(n\times 1)} \quad . \tag{1.1.2}$$

Let us write assumption (3) as $\underset{\sim}{Z} \sim (\underset{\sim}{0},\sigma^2 I)$, or, more strongly, $\underset{\sim}{Z} \sim N(0,\sigma^2 I)$. Then $E(\underset{\sim}{Y}) = X\underset{\sim}{\beta}$ and $E\{(\underset{\sim}{Y}-X\underset{\sim}{\beta})(\underset{\sim}{Y}-X\underset{\sim}{\beta})'\} = \sigma^2 I$. Applying the method of maximum likelihood to the density of $\underset{\sim}{Y}$, the least squares estimates of the unknown model parameters $(\underset{\sim}{\beta})$ are obtained and are given by the well-known formula $\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{Y}$. Under the assumption that $\underset{\sim}{Z} \sim N(0,\sigma^2 I)$, the vector $\hat{\underset{\sim}{\beta}}$ has the following properties:

(a) $E(\hat{\underset{\sim}{\beta}}) = \underset{\sim}{\beta}$

(b) $E\{(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})'\} = \sigma^2 (X'X)^{-1}$

(c) $\hat{\underset{\sim}{\beta}}$ is the BLUE (best linear unbiased estimator) for $\underset{\sim}{\beta}$ (Gauss-Markov Theorem).

(d) $\hat{\underset{\sim}{\beta}}$ has minimum variance among all LUE for $\underset{\sim}{\beta}$ (Gauss-Markov Theorem).

(e) $\hat{\underset{\sim}{\beta}}$ is a maximum likelihood estimate.

(f) $\hat{\underset{\sim}{\beta}} \sim N\{\beta, \sigma^2(X'X)^{-1}\}$

Suppose now, that assumption (3) is violated and that $\underset{\sim}{Z} \sim N\{0,\Sigma\}$, where $\Sigma = \sigma^2\Psi$, Ψ a general covariance matrix. Then $E(\underset{\sim}{Y}) = X\underset{\sim}{\beta}$ and $E\{(\underset{\sim}{Y}-X\underset{\sim}{\beta})(\underset{\sim}{Y}-X\underset{\sim}{\beta})'\} = \Sigma = \sigma^2\Psi$. In order to obtain estimates of the elements of $\underset{\sim}{\beta}$, the maximum likelihood procedure is again applied to the density of Y. This yields the estimator $\hat{\underset{\sim}{\beta}} = (X'\Psi^{-1}X)^{-1}X'\Psi^{-1}\underset{\sim}{Y}$, sometimes called the Markov estimator. From the Gauss-Markov Theorem, it can be shown that $\hat{\underset{\sim}{\beta}}$ is the BLUE for $\underset{\sim}{\beta}$ and that $\hat{\underset{\sim}{\beta}}$ has minimum variance $(\sigma^2(X'\Psi^{-1}X)^{-1})$ among all LUE for $\underset{\sim}{\beta}$. Let us consider the plight of the researcher who (in place of the more general Markov estimator) utilizes the usual least squares estimate of $\underset{\sim}{\beta}$, say $\hat{\beta}_L$, in the presence of dependent errors. Only if the characteristic vectors of Ψ have a specific form, does $\hat{\underset{\sim}{\beta}}_L$ equal $\hat{\underset{\sim}{\beta}}$ [4]. However, this is usually not the case in any practical situation. Assuming that the two estimators are unequal, the vector $\hat{\underset{\sim}{\beta}}_L$ has the following properties:

(a) $E(\hat{\underset{\sim}{\beta}}_L) = \underset{\sim}{\beta}$

(b) $E\{(\hat{\underset{\sim}{\beta}}_L-\underset{\sim}{\beta})(\hat{\underset{\sim}{\beta}}_L-\underset{\sim}{\beta})'\} = (X'X)^{-1}X'\Sigma X(X'X)^{-1} = \sigma^2(X'X)^{-1}X'\Psi X(X'X)^{-1}$

(c) $\hat{\underset{\sim}{\beta}}_L$ no longer has minimum variance among all LUE for $\underset{\sim}{\beta}$

(d) $\hat{\underset{\sim}{\beta}}_L$ is no longer a maximum likelihood estimator

(e) $\hat{\underset{\sim}{\beta}}_L \sim N\{\underset{\sim}{\beta}, \sigma^2(X'X)^{-1}X'\Psi X(X'X)^{-1}\}$

The implication of properties (b) and (e) is that the usual least squares formula for finding the variance of $\hat{\underset{\sim}{\beta}}_L$, $\sigma^2(X'X)^{-1}$, yields a biased estimate of the true variance $E\{(\hat{\underset{\sim}{\beta}}_L - \underset{\sim}{\beta})'(\hat{\underset{\sim}{\beta}}_L - \underset{\sim}{\beta})\} = \sigma^2(X'X)^{-1}X' X(X'X)^{-1}$. The least squares test statistics used for making inferences about the elements of $\underset{\sim}{\beta}$ are often stochastically larger in the presence of dependent errors. Thus, the researcher will tend to reject the null hypothesis $H_o: \beta_i = 0$ more often than he should, and hence, will include insignificant terms in the model.

In order to avoid the problems introduced by applying the least squares procedure when the errors are dependent, a test for independence in the errors (sometimes called a test for serial correlation of the errors) is essential.

1.2 Various Tests

Since the errors ($\underset{\sim}{Z}$) of a least squares regression are not observable, a test for serial correlation of the errors must be based on the residuals ($\hat{\underset{\sim}{Z}}$) obtained from the calculated regression, where $\hat{\underset{\sim}{Z}} = \underset{\sim}{Y} - X\hat{\underset{\sim}{\beta}}$. However, in most cases the residuals are correlated whether the errors are dependent or not; consequently, ordinary tests of independence for an observed sequence of random variables cannot be used without modification. Many test procedures have been proposed for this problem, and in this section a summary of several popular tests will be presented.

1.2.1 The Durbin-Watson Test

J. Durbin and G.S. Watson [10] in 1950-51 were among the first to consider the problem of testing for serial correlation of the errors in least squares regression, and their test and its modified versions remain the most widely applied tests of regression error independence. In order to investigate the random error component $\underset{\sim}{Z}$ in the model

$$\underset{\sim}{Y}_{(n\times 1)} = X_{(n\times g)}\underset{\sim}{\beta}_{(g\times 1)} + \underset{\sim}{Z}_{(n\times 1)}, \qquad (1.2.1)$$

Durbin-Watson (D-W) make the usual assumption under the null hypothesis: the errors are independent and identically distributed normal random variables with zero mean and constant variance, i.e., $Z_t \sim$ i.i.d $N(o,\sigma^2)$. The type of alternative they consider is one for which the vector of random errors, $\underset{\sim}{Z}$, is stationary, and for which $\mathrm{Cov}\{Z_t, Z_{t+h}\} = \gamma(h)$ is exponentially (or geometrically) decreasing in h. A convenient model for this hypothesis is the first order autoregressive [AR(1)] model, or as it is sometimes called, the stationary Markov process: $Z_t = \emptyset Z_{t-1} + \varepsilon_t$, where $|\emptyset| < 1$ and $\varepsilon_t \sim$ i.i.d $N(o,\sigma^2)$ independently of the Z_t's. Thus, the null and alternative hypotheses take the form:

$$H_o\colon \emptyset = 0 \quad \text{versus} \quad H_a\colon \emptyset \neq 0 \quad .$$

The test statistic derived by (D-W) is given by:

$$d = \frac{\sum_{t=2}^{n} (\hat{Z}_t - \hat{Z}_{t-1})^2}{\sum_{t=1}^{n} \hat{Z}_t^2} \quad ,$$

where $\underset{\sim}{\hat{Z}}$ is the vector of residuals from the calculated least squares regression $(\underset{\sim}{\hat{Z}} = \underset{\sim}{Y} - X\underset{\sim}{\hat{\beta}})$.

Durbin and Watson derived the d statistic using a result obtained by T.W. Anderson [5] in 1948. Anderson showed that no test exists which is uniformly most powerful against the two-sided alternative. However, for cases in which the columns of the X (design) matrix in the model (called 'regression vectors' by Anderson) are linear combinations of the eigenvectors (latent vectors) of a matrix A (where A occurs in the density function of $\underset{\sim}{Z}$), tests based on $\frac{\underset{\sim}{\hat{Z}}' A \underset{\sim}{\hat{Z}}}{\hat{Z}'\hat{Z}}$ are uniformly most powerful (UMP) against one-sided alternatives, and have optimal properties for two-sided alternatives. The density functions of $\underset{\sim}{Z}$ which lead to the UMP test have

have been shown by Anderson to take the form:

$$f(\underset{\sim}{Z}:\phi,\sigma^2) = K \exp[\{(1+\phi)^2\underset{\sim}{Z}'\underset{\sim}{Z} - 2\phi\underset{\sim}{Z}'A\,\underset{\sim}{Z}\} \cdot (-\frac{1}{2\sigma^2})]. \qquad (1.2.2)$$

When the distribution of $\underset{\sim}{Z}$ takes the form of the AR(1) model, the density of $\underset{\sim}{Z}$ can be written as:

$$f(\underset{\sim}{Z}:\phi,\sigma^2)= K_2\exp[-\frac{1}{2\sigma^2}\{(1+\phi)^2\sum_{t=1}^{n} Z_t^{\,2} - \phi^2(Z_1^{\,2}+Z_n^{\,2})$$
$$- 2\phi\sum_{t=2}^{n} Z_tZ_{t-1}\}] . \qquad (1.2.3)$$

By taking $\underset{\sim}{Z}'A\,\underset{\sim}{Z} = \sum_{t=2}^{n}(Z_t-Z_{t-1})^2$ in equation (1.2.2), (D-W) obtained the density function:

$$f(\underset{\sim}{Z}:\phi,\sigma^2) = K_3\exp[-\frac{1}{2\sigma^2}\{(1+\phi)^2\sum_{t=1}^{n} Z_t^{\,2} - \phi(Z_1^{\,2}+ Z_n^{\,2})$$
$$- 2\phi\sum_{t=2}^{n} Z_tZ_{t-1}\}] , \qquad (1.2.4)$$

which is very close to the AR(1) density in equation (1.2.3). Thus, an appropriate choice of A, namely

$$A = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots\; 2 & -1 \\ 0 & 0 & 0 & \cdots\; -1 & 1 \end{bmatrix} ,$$

leads to a d statistic which provides a UMP test against one-sided alternatives in the "latent vector case", i.e., the case when the columns of the design matrix coincide with the latent vectors of A.

The d statistic is bounded below by 0 (obvious) and above by 4. In order to see that $d \leq 4$, expand the numerator of the statistic:

$$\sum_{t=2}^{n}(Z_t-Z_{t-1})^2 = \sum_{t=2}^{n}(Z_t^2-2Z_tZ_{t-1}+Z_{t-1}^2)$$

$$= \sum_{t=2}^{n} Z_t^2-2\sum_{t=2}^{n} Z_tZ_{t-1}+\sum_{t=2}^{n} Z_{t-1}^2$$

$$\leq \sum_{t=1}^{n} Z_t^2-2\sum_{t=2}^{n} Z_tZ_{t-1}+\sum_{t=1}^{n} Z_t^2$$

Then

$$d = \frac{\sum_{t=2}^{n}(Z_t-Z_{t-1})^2}{\sum_{t=1}^{n} Z_t^2} \leq 2 - \frac{2\sum_{t=2}^{n} Z_tZ_{t-1}}{\sum_{t=1}^{n} Z_t^2} = 2\left[1 - \frac{\sum_{t=2}^{n} Z_tZ_{t-1}}{\sum_{t=1}^{n} Z_t^2}\right]$$

Now,

$$\frac{\sum_{t=2}^{n} Z_tZ_{t-1}}{\sum_{t=1}^{n} Z_t^2} \geq -1 ,$$

hence

$$d \leq 2[1-(-1)] = 4 .$$

Thus, if the errors were positively serially correlated, e.g., $Z_t = \emptyset Z_{t-1} + \varepsilon_t$, $\emptyset > 0$, d would tend to be relatively small, and if the errors were negatively serially correlated, e.g., $Z_t = \emptyset Z_{t-1} + \varepsilon_t$, $\emptyset < 0$, d would tend to be relatively large. The user of the d test interested in detecting the existence of positive serial correlation of the regression errors would reject the null hypothesis of independent errors ($\emptyset = 0$) in favor of the alternative $\emptyset > 0$ if $d < d^*$, where d^* is the appropriate critical value of d for a lower-tailed test. Upper-tailed tests (tests for negative serial correlation) are carried out in a similar manner.

A major drawback to this testing procedure is that it is only possible to find the exact null distribution of d in very special cases (see Anderson [5]), cases which do not frequently occur in practice. However, D-W show that it is possible to calculate upper (d_U*) and lower (d_L*) bounds to the critical value d*, where $0 \leq d_L^* < d^* < d_U^* \leq 4$. Tabulated values of the critical bounds are made available in Durbin and Watson [10]. The procedure for conducting a test designed to detect the existence of positive serial correlation is as follows:

(a) if $d < d_L^*$, reject H_o (significant at level α)

(b) if $d > d_U^*$, do not reject H_o (insignificant at level α)

(c) if $d_L^* < d < d_U^*$, the test is inconclusive.

By observing that the null distribution of d is symmetric, with $0 \leq d \leq 4$, a test for negative serial correlation can be carried out exactly as above, only using the statistic 4-d. Two-sided tests are obtained by combining single tail areas, using $\alpha/2$.

When the observed value of d falls in the "region of ignorance", $d_L^* < d < d_U^*$, the bounds test is inconclusive. It is here that D-W recommend applying a beta approximation. The procedure involves transforming d so that its range is (0,1), accomplished by the simple transformation $\frac{1}{4}d$. A beta distribution with the same mean and variance of d is then fit using the method of moments. Critical values of $\frac{1}{4}d$ are found from Tables of the Incomplete Beta Function, and the test is conducted in the usual manner.

Tests for the case in which the errors in a least squares regression follow a specific non-first order autoregressive process were developed by Peter Schmidt [29] and K.F. Wallis [34] and were based upon the work done by D-W. Schmidt considered the AR(2) alternative: $Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \varepsilon_t$, $\varepsilon_t \sim$ i.i.d(o,σ^2). A test of the null hypothesis H_o: $\emptyset_1 = \emptyset_2 = 0$ is carried out by considering a generalization of the D-W d statistic:

$$d_2 = \frac{\sum_{t=2}^{n}(\hat{Z}_t-\hat{Z}_{t-1})^2 + \sum_{t=3}^{n}(\hat{Z}_t-\hat{Z}_{t-3})^2}{\sum_{t=1}^{n}\hat{Z}_t^2}.$$

When analyzing quarterly data, Wallis argued that researchers may be interested in detecting the existence of seasonal variation of the error terms. One possible seasonal error model is the AR(4) process: $Z_t = \emptyset Z_{t-4} + \varepsilon_t$, $\varepsilon_t \sim \text{i.i.d}(o,\sigma^2)$. Wallis' "seasonal effects" test is based upon the value of the modified D-W statistic:

$$d_4 = \frac{\sum_{t=5}^{n}(\hat{Z}_t-\hat{Z}_{t-4})^2}{\sum_{t=1}^{n}\hat{Z}_t^2} \quad .$$

Critical bounds for both non-first order tests have been tabulated by their respective developers. The tests are conducted in exactly the same manner as the D-W bounds test.

1.2.2 Beta Approximation

H. Theil and A.L. Nagar [33] criticized the D-W bounds test because of its inconclusive region (d_L*, d_U*), and fear that the practical research worker may interpret "no inference possible" as equivalent with "no evidence to reject the null hypothesis of independence." This results in bias in the sense that many cases of positive serial correlation will be overlooked, especially when the inconclusive region is large. This situation occurs when the number of observations is small and the number of independent variables in the model is large.

The Theil-Nagar (T-N) approach is to fit a beta distribution immediately, dispensing with finding bounds for the critical value of d*. However, their method of fitting the beta distribution differs from that used by D-W. The d statistic is transformed to a beta (p,q) variable in the range (0,1) by considering the quantity $X = \frac{d-a}{4-(a+b)}$, where the range of d is (a,b). The parameters p, q, a, and b are then determined by the methods of moments, using the first four moments of d. The first two moments of d are used to find expressions for a and b, while the third and fourth moments of d are utilized for determining p and q. T-N simplify the necessary calculations considerably by ignoring terms of lower order in the moment expressions.

R.C. Henshaw [20] also considered a beta approximation to the distribution of d. However, where T-N approximate the first four moments of d, Henshaw uses the exact moments of d in fitting the beta distribution. The transformation $X = \frac{d-\nu_1}{\nu_{n-g}-\nu_1}$ is used, where ν_1 and ν_{n-g} are, respectively, the smallest and largest latent roots of $(I-X(X'X)^{-1}X')A$. The parameters p, q, ν_1, ν_{n-g} are then found by the method of moments.

The calculations needed in Henshaw's test are much more extensive than those of T-N. However, T-N point out that their approximations are good only when "the behavior of the [independent] variables is sufficiently smooth in the sense that the first and second differences are small compared with the range of the . . . variable itself."

1.2.3 d_U Approximation

For polynomial regression, E.J. Hannan [17] discovered that the D-W upper bounding statistic, d_U, gives a very good approximation to the true distribution of d. Hannan's reasoning was based on the fact that the d_U bound is attained in the latent vector case, i.e., when the g columns of the design matrix are linear combinations of the g latent vectors vectors associated with the smallest non-zero latent roots of the matrix A. Later, Hannan and R.D. Terrell [18] showed that the approximation remains reasonably good whenever the coded values of the independent variables are concentrated near the origin (spectrum of the independent variables in the model is concentrated near the origin), which frequently occurs for economic time series.

In the light of Hannan's work, Durbin and Watson [12] theorized that the shape of the d distribution might well be better approximated by the distribution of the upper bounding statistic d_U. However, D-W felt that this approximate distribution should have the same mean and variance as the distribution of d since this was the one desirable property possessed by the beta approximation. The statistic $d^* = a + bd_U$ was considered; where a and b are chosen so that d^* has the same first two moments as d. Critical values of d_U, tabulated by D-W in their 1951 paper, are used in finding the critical values of d^* [11].

1.2.4 BLUS Estimators

It can be shown that the covariance matrix for the residuals is

$$E(\underset{\sim}{\hat{Z}}\, \underset{\sim}{\hat{Z}}') = \sigma^2\{I-X(X'X)^{-1}X'\} = \sigma^2 M.$$

Thus, even if the null hypothesis is true, the residuals are, in general, correlated and heteroscadastic. The result complicates matters considerably, as evidenced by D-W's upper and lower bounds to the significant points of d. H. Theil [32] reasoned that the "testing procedure would be simplified considerably if the covariance matrix of the [residuals] were of the form $\sigma^2 I$ rather than $\sigma^2 M$." Theil considered new estimates of the regression errors, called BLUS estimates; that is, best linear unbiased within the class of those linear unbiased estimates that have scaler covariance matrix $\sigma^2 I$.

The new regression residuals, $\underset{\sim}{V}$, are formed such that $E\{(\underset{\sim}{V}-\underset{\sim}{Z})'(\underset{\sim}{V}-\underset{\sim}{Z})\}$ is minimized subject to $E(\underset{\sim}{V}) = E(\underset{\sim}{Z}) = 0$ and $E(\underset{\sim}{V}\underset{\sim}{V}') = \sigma^2\Omega$, where Ω is chosen apriori. By choosing $\Omega_{(nxn)}$ as a diagonal matrix with n-g ones and g zeroes along the diagonal, the new residuals $\underset{\sim}{V}$ are BLUS estimators of $\underset{\sim}{Z}$. Theil's test statistic takes the form of D-W's d statistic, only with the $\hat{Z}_t$'s replaced by the V_t's:

$$d_{BLUS} = \frac{\sum_{t=2}^{n} (V_t - V_{t-1})^2}{\sum_{t=1}^{n} V_t^2} \quad .$$

Now the BLUS estimates, under the null hypothesis, are independent with zero mean and constant variance, hence, the exact null distribution of d_{BLUS} is a standard one. Critical points are tabulated by B.I. Hart [19].

Theil notes that the choice of the matrix Ω actually results in a dropping of g of the n residuals from the estimation process. This procedure permits any set of g residuals to be dropped, the choice being left to the user. However, as pointed out by A.P.J. Abrahamse and A.S. Louter [2], there exist $\binom{n}{g}$ possible BLUS estimators, each of which will, in general, yield a different result. No method has been given for selecting the optimal vector of BLUS estimators.

Abrahamse and Louter (A-L) hypothesize that a diagonal matrix Ω is not a necessary condition to eliminate the inconclusive region in the D-W bounds test. It is sufficient to require that Ω does not depend on the design matrix. The estimators of A-L are derived in the same manner as the BLUS estimators with the modification that $\sigma^2\Omega = \sigma^2 KK'$ approximates $\sigma^2 M$, and the test statistic, like Theil's, takes the form:

$$Q = \frac{\sum_{t=2}^{n} (V_t - V_{t-1})^2}{\sum_{t=1}^{n} V_t^2} \quad .$$

A-L show that the Q statistic has the same distribution as the D-W upper bounding statistic d_U. Hence, it is an exact test, with tables available for finding significance points.

The procedures developed by Theil and A-L both necessitate tedious calculations in order to obtain residual estimates. G.D.A. Phillips and A.C. Harvey [26] sought to simplify the calculations and interpretations of the estimated residuals, while retaining the properties of the BLUS estimates. They determined that this could be accomplished by utilizing "recursive residuals." One normally estimates the model parameters on the basis of all n observations. The recursive residuals, denoted by $\tilde{Z}_j$, are found by using only the first j observations, j=g+1, g+2, . . ., n. These residuals can be easily generated by a recursive algorithm. The Phillips and Harvey (P-H) test statistic takes the familiar form:

$$\tilde{d} = \frac{\sum_{t=g+2}^{n} (\tilde{Z}_t - \tilde{Z}_{t-1})^2}{\sum_{t=1}^{n} \tilde{Z}_t^2} .$$

Under H_0, the $\{\tilde{Z}_t\}$ are distributed independently with zero mean and constant variance. Critical values of this standard distribution have been tabulated by Hart [19]. Like Theil's BLUS estimators, the recursive residuals are linear, unbiased, and possess a scaler covariance matrix.

1.2.5 The Sign Test

One of the first researchers to consider a non-parametric alternative to the D-Q bounds test was R.C. Geary [14]. Geary found that a simple count of the number of sign changes in the least squares residuals could be used as a test for serial correlation in the errors. If the errors are independent, the signs (plus or minus) of the residuals in the sequence will occur in random order. Serial correlation is inferred then, if the number of sign changes are fewer than expected.

Let n be the number of observations and τ the number of sign changes in the sequence of residuals. Then under the null hypothesis of independence, τ has the distribution:

$$P(\tau) = \frac{(n-1)!}{\tau!(n-1-\tau)!}$$

Thus, it is possible to calculate significance points of τ under H_0. The hypothesis of white noise is rejected when $\tau < \tau^*$, where τ^* is the tabulated significance point for the values of n and α.

Geary notes that the τ-test is very similar to the familiar runs test. In fact, the number of runs, U, is one more than the number of sign changes, τ, i.e., $U = \tau + 1$. In runs theory, the number of plus and minus signs are taken into account, whereas these are ignored in the sign test. Geary remarks that the sign test is a handy tool for assessing the probable presence of serial correlation for those workers in multivariate regression time series who are computerless, or without the Durbin-Watson routine in their computer.

1.2.6 Periodogram Based Tests

Recall that D-W's 1950 d test was designed to have high power against only a first order autoregressive alternative error model. However, the nature of the residual dependence, when it exists, may be more general (non-first order). Thus, an investigator may wish to obtain a more comprehensive picture of the departure from serial independence than is provided by a single statistic like d. Durbin [9] suggests that it may be more appropriate to ask what the data reveals about the departure from serial independence rather than to set up a particular parametric alternative and seek a test which has high power against it. Durbin's 1969 technique for studying the general nature of the serial dependence in a stationary series of observations $Z_1, Z_2, \ldots, Z_T$ is to compute the quantities

$$S_j = \sum_{r=1}^{j} P_r / \sum_{r=1}^{m} P_r \quad , \; r = 1, 2, 3, \ldots, m$$

where

$$P_r = \frac{2}{T} \left| \sum_{t=1}^{T} Z_t e^{(2\pi irt)/T} \right|^2 ,$$

and then to make a plot of S_j versus j/m. This plot is called the cumulated periodogram, S_j is its sample path, and r = 1, 2, . . ., m the periodogram ordinates.

It is known that when $Z_1, Z_2, \ldots, Z_T$ are independent and identically distributed, the sample path S_j behaves asymptotically like that of the sample distribution function in the theory of order statistics. Hence, the Kolmogorov-Smirnov limits can be applied to provide a test of serial independence. However, when the periodogram is computed from the least squares residuals (Z_t), modifications to this procedure are necessary, since the residuals themselves are correlated in any case.

For a test against positive serial correlation, a critical value C_o for the Kolmogorov-Smirnov statistic is acquired from available tables. Like the D-W bounds procedure, there are three regions for the sample path of S_j:

(1) If the sample path of S_j crosses an upper line $C_o + j/m'$, reject H_o.

(2) If the sample path of S_j fails to cross a lower line $C_o + [j - \frac{1}{2}(g-1)]/m'$, do not reject H_o (where $\frac{1}{2}g - \frac{1}{2} \leq j \leq m$ and $m' = \frac{1}{2}(n-g)$).

(3) If the sample path crosses neither the lower or upper line, the test is inconclusive (analagous to the inconclusive region for the D-W bounds test).

Durbin notes that this technique is conservative in the sense that, for a test of significance level α, the probability of falsely rejecting the null hypothesis does not exceed α. Tests against negative serial correlation are carried out similarly, and a two-sided test may be obtained by applying both one-sided regions at significance level $\alpha/2$.

1.3 Concluding Note

We have presented a brief summary of available tests for serial correlation of the errors in a regression analysis. The properties of these tests are discussed, and comparisons made, in the chapter that follows. In particular, we are concerned with the performance of these tests in the presence of non-first order serially correlated errors.

CHAPTER II

COMPARING TESTS OF SERIAL CORRELATION

2.1 Power Comparisons

Many researchers have conducted Monte Carlo and simulation studies in order to estimate the powers of the various tests for serial independence. Most of the power studies in the literature focus on the Durbin-Watson d test. In the following sections, the d test is compared to three groups of alternative tests and a brief summary of the results is presented.

2.1.1 Exact d Test Versus Approximate Tests of Serial Correlation

In 1971, Durbin and Watson [12] investigated and compared the exact distribution of d with certain approximate tests which were developed using their procedure as a basis. These include the Durbin-Watson, Theil-Nagar, and Henshaw beta approximations and the Hannan d_U and Durbin-Watson $a+bd_U$ approximations.

Letting d' denote a random variable having a distribution used as an approximation to the true distribution of d, D-W computed $P(d' \leq d_\alpha)$ for each approximating distribution, where $P(d \leq d_\alpha) = \alpha$. The power comparisons reveal that the tests fall into three categories: Theil-Nagar beta and Hannon d_U; D-W beta and D-W $a+bd_U$; Henshaw beta. These groups are given in increasing order of accuracy, which coincides with increasing difficulty of calculation.

As a result of their research, D-W make the following recommendation: when a test for serial correlation based on the least squares residuals is required, the D-W d test should be applied first. If the result is inconclusive, apply the D-W beta or $a+bd_U$ approximation. When special accuracy is needed, use the Henshaw beta approximation. Finally, if a more comprehensive picture of the serial properties of the errors is desired than is provided by the value of a single statistic, employ the cumulated periodogram method.

2.1.2 d Test Versus BLUS Test

Abrahamse and Koerts [1] calculated the power of the BLUS test for various values of n and g (number of independent variables) with the parameter of the alternative error hypothesis at a relatively large value of $\emptyset = .8$. This was compared with the following three quantities:

(a) $P(d<d_L^*|H_a)$, called the probability of a correct decision for the d test, where d_L^* is the lower critical bound for d,

(b) $P(d>d_U^*|H_a)$, called the probability of an incorrect decision for the d test, where d_U^* is the upper critical bound for d,

(c) $P(d<d^*|H_a)$, called the power of the d test, where d^* is the approximate true significance point of d obtained by using a method similar to those discussed in Chapter I.

Note that the probability of a correct decision, as defined by Abrahamse and Koerts, is the power of the d test when using D-W's bounds.

The results obtained by A-K can be summarized as follows:

(i) The power of the d test generally exceeds the power of the BLUS test.

(ii) The power of the BLUS test dominates the probability of a correct decision for the d test. The smaller the number of degrees of freedom (n-g), the greater the difference.

(iii) The power of the BLUS test, the probability of a correct decision and the power of the d test converge as n-g increases, for in this case the inconclusive region of the d test disappears.

Since calculation of approximate significance points for the d test may require many more computations than the BLUS procedure, A-K recommend one follow the schematic shown in Figure 2.1.1 when choosing a test for residual correlation.

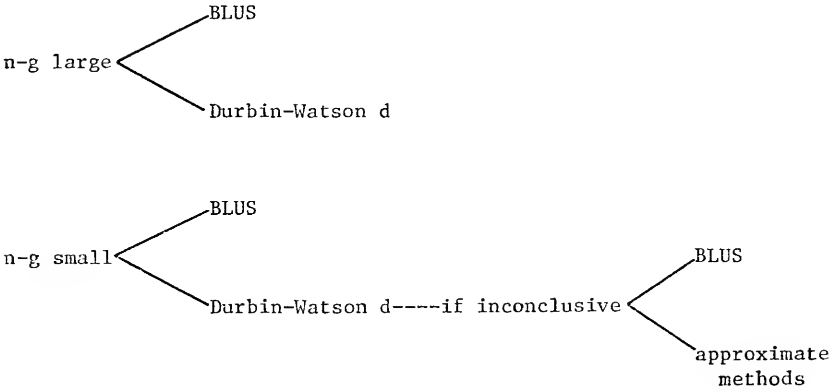


FIGURE 2.1.1

ABRAHAMSE-KOERTS CHOICE OF A TEST FOR SERIAL CORRELATION

2.1.3 The Durbin-Watson d Test Versus Non-First Order Alternative Tests

V. Kerry Smith [30] compared the approximate power of four tests for serial correlation with two sets of alternatives: second order autoregressive processes, denoted AR(2), and first order moving average processes, denoted MA(1). The four tests considered in the study were the D-W d test, the Durbin cumulated periodogram (S_j) test, the Schmidt (dz) test and the Geary (τ) count of signs change test. Smith's Monte Carlo results are summarized as follows:

(a) For the MA(1) error model, the d test proved to be consistently more powerful than the other procedures. However, with small samples and low first order correlation coefficients, the differences in the tests are not pronounced.

(b) For the AR(2) error model, three main results are obtained.

  (i) For certain values of ϕ_1 and ϕ_2 in the AR(2) model, the d test is more powerful than the d_2 test. The τ test is shown to be much inferior to the d.

  (ii) In models with large and approximately equal values of ϕ_1 and ϕ_2 the d_2 test appears to become more powerful than the others. (Large ϕ_1 and ϕ_2 imply first and second order autocorrelations near 1. For a detailed discussion of autocorrelation, see Section 3.3.1.)

  (iii) The S_j test appears to be at least as powerful as the d test for most of the values of ϕ_1 and ϕ_2 considered by Smith. However, the S_j test does not perform as well as the d_2 test in most instances.

2.2 The Alternative Error Model

Almost all of the tests discussed in Section 1.2 consider only a first order autoregressive hypothesis, i.e., one in which the hypothesized error model under H_a takes the form

$$Z_t = \phi Z_{t-1} + \varepsilon_t,$$

where $\{\varepsilon_t\} \sim$ i.i.d. $(0,\sigma^2)$ independently of $\{Z_t\}$. The primary reason for this restrictive hypothesis is that the distributions of the test statistics are complex, even in this simple case. However, many time series exhibit autocorrelation patterns which are not well modeled by the first order model. When analyzing quarterly residuals, it may be more appropriate to consider the model

$$Z_t = \phi Z_{t-4} + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is defined as above. The performance of the Durbin-Watson d or similar test should be evaluated for models like these.

In the sections that follow, we will examine the properties of tests designed for the general alternative autoregressive-moving average (ARMA) model

$$Z_t + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ . \ . \ . \ \phi_p Z_{t-p} = \varepsilon_t + \theta_1 {}_{t-1}$$

$$+ \theta_2 {}_{t-2} + \ . \ . \ . + \theta_q \varepsilon_{t-q} .$$

CHAPTER III

TWO TESTS DESIGNED TO DETECT A GENERAL ALTERNATIVE

3.1 Box and Pierce "Portmanteau" Test

G.E.P. Box and D.A. Pierce [7] developed a method for checking the adequacy of fit in autoregressive-moving average (ARMA) time series models. Finding a test for serial independence of the errors in least squares regression was not of prime importance to Box and Pierce (B-P). However, they did discover that the statistic used for checking model adequacy is, under certain conditions, very nearly equivalent to the D-W d test.

Consider the general ARMA (p,q) model:

$$Z_t = \sum_{j=1}^{p} \emptyset_j Z_{t-j} + \sum_{i=1}^{q} \theta_i a_{t-i} + a_t, \quad t = 1,2,. . ., n \qquad (3.1.1)$$

where the $\{a_t\}$ are independent unobservable random variables, and the vector of model parameters $(\emptyset_1, \emptyset_2, . . ., \emptyset_p, \theta_1, \theta_2, . . ., \theta_q)'$ is unknown. In order for (3.1.1) to attain stationarity, the autoregressive parameters $(\emptyset_1, \emptyset_2, . . ., \emptyset_p)$ must satisfy certain constraints. Fuller [13], and others show that the condition of stationarity is met if the roots of the equation $x^p + \emptyset_1 x^{p-1} + \emptyset_2 x^{p-2} + . . . + \emptyset_p = 0$ are all in the unit circle. To guarantee the invertibility of the moving average part of model (3.1.1), Fuller states that the roots of the equation $y^q + \theta_1 y^{q-1} + \theta_2 y^{q-2} + . . . + \theta_q = 0$ must also lie in the unit circle. This condition on $(\theta_1, \theta_2, . . ., \theta_q)$ enables the researcher to express the $\{Z_t\}$ generated from the MA process,

$$Z_t = \sum_{i=0}^{q} \theta_i a_{t-i},$$

as an AR time series,

$$Z_t = \sum_{j=1}^{\infty} \phi_j Z_{t-j} + a_t \quad .$$

(We utilize this invertibility property later in Section 3.3.) Throughout, then, we assume that model (3.1.1) is stationary and invertible, i.e., the roots of the equations

$$x^p + \phi_1 x^{p-1} + \phi_2 x^{p-2} + . . . + \phi_p = 0$$

and

$$y^q + \theta_1 y^{q-1} + \theta_2 y^{q-2} + . . . + \theta_q = 0$$

are all in the unit circle.

After calculating the parameter estimates, $\hat{\phi}_1$, $\hat{\phi}_2$, . . ., $\hat{\phi}_p$, $\hat{\theta}_1$, $\hat{\theta}_2$, . . ., $\hat{\theta}_q$, by the usual time series methods (see Box-Jenkins [8]), the researcher may then be interested in checking the adequacy of the model:

$$Z_t = \sum_{j=1}^{p} \hat{\phi}_j Z_{t-j} + \sum_{i=1}^{q} \hat{\theta}_i a_{t-1} + a_t \quad . \tag{3.1.2}$$

B-P proposed that the first K sample autocorrelations of the residual series $\{a_t\}$, r_1, r_2, . . ., r_K, be used for this purpose, where K is small relative to n, and where

$$r_j = \frac{\sum_{t=j+1}^{n} a_t a_{t-j}}{\sum_{t=1}^{n} a_t^2} , \quad j = 1, 2, . . ., K . \tag{3.1.3}$$

In particular, B-P considered the statistic,

$$n \sum_{j=1}^{K} r_j^2 \quad .$$

However, the true sample autocorrelations cannot be calculated since the $\{a_t\}$ in (3.1.1) are unobservable. In order to proceed, the researcher may apply the method of least squares to (3.1.2), to obtain the residuals $\{\hat{a}_t\}$. Replacing the $\{a_t\}$ with $\{\hat{a}_t\}$ in (3.1.3), estimates of the first K sample autocorrelations are found, i.e.,

$$\hat{r}_j = \frac{\sum_{t=j+1}^{n} \hat{a}_t \hat{a}_{t-j}}{\sum_{t=1}^{n} \hat{a}_t^2}, \qquad j = 1, 2, \ldots, k. \tag{3.1.4}$$

The statistic $Q = n \sum_{j=1}^{K} \hat{r}_j^2$ is then of interest.

B-P shows that Q will, to a close approximation, have an asymptotic χ^2-distribution with K-p-q degrees of freedom. The 'portmanteau' test for checking the adequacy of any ARMA (p,q) process will lead to a rejection of the hypothesis that the model adequately fits the data if Q is "too large", where $Q \sim \chi^2_{(K-p-q)}$.

Although B-P are not actually considering a regression model in which the errors are to be tested for independence, notice that if p=q=0 in (3.1.1), i.e., if the $\{Z_t\}$ are assumed independent, a test based on

$$r_1 = \left(\sum_{t=2}^{n} \hat{Z}_t \hat{Z}_{t-1}\right) / \left(\sum_{t=1}^{n} \hat{Z}_t^2\right),$$

is very nearly equivalent to the D-W d test. This fact would lead one to believe that the 'portmanteau' test may have applications in the area of testing for serial independence of the errors in a regression analysis. In particular, let $p \neq 0$, $q = 0$, in (3.1.1). The model then takes the form:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t \ , \qquad (3.1.5)$$

where $\{\varepsilon_t\}$ are independently distributed. If we let $\{Z_t\}$ represent the unobservable errors in a regression model, then (3.1.5) is the general autoregressive alternative error model discussed in Section 2.2. Under the null hypothesis of white noise, $p = 0$, and the B-P statistic Q is distributed asymptotically as a chi-square random variable with K degrees of freedom. It seems reasonable to utilize the Q statistic in a test for serial correlation of the $\{Z_t\}$. "Large" values of Q would lead to a rejection of the hypothesis of white noise. Further, since the form of (3.1.5) is more general than the alternative model considered by D-W and others (see Section 1.2), one would expect the 'portmanteau' test to be more powerful against general alternatives.

3.2 A New Technique: The Max-X^2 Procedure

James T. McClave [24] considered a new method of estimating the order of autoregressive time series models, called the maximum chi-square (max-X^2) technique. As was the case with B-P, a test for serial correlation of the errors in a regression model was not of prime importance in McClave's research. However, after developing the max-X^2 procedure, McClave hypothesized that this technique could also be used as a test for serial correlation of the regression errors, and that this test would have high power against general alternatives.

McClave considered the stationary autoregressive (AR) model,

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_m Z_{t-m} + \varepsilon_t \ , \qquad (3.2.1)$$

where the $\{\varepsilon_t\}$is a white noise process. Call model (3.2.1) a p^{th} order subset AR model with maximum lag m if exactly p of the lag coefficients

$\emptyset_1, \emptyset_2, \ldots, \emptyset_m$ are nonzero, $\emptyset_m \neq 0$ and $\emptyset_j = 0, \forall_j > m$. The researcher whose ultimate goal is to obtain estimates of the nonzero lag coefficients in (3.2.1), must first estimate the elements of the parameter set $\textcircled{H} = \{p;\ j_1, j_2, \ldots, j_p\}$, where $(j_1, j_2, \ldots, j_p)$ are the lags having nonzero coefficients. The initial step in the order estimation process is the application of the 'subset autoregression algorithm' developed by McClave [24]. Subset autoregression enables the researcher to generate a sequence of estimates of $\textcircled{H}$, say $S = \{\hat{\textcircled{H}}_1, \hat{\textcircled{H}}_2, \ldots, \hat{\textcircled{H}}_K\}$, where $\textcircled{H}_k = \{k;\ j_1, j_2, \ldots, j_k\}$. Once these estimates are obtained, McClave proposes the following sequential procedure for choosing the order from S:

1. Compute the sequence of statistics $\{\hat{M}_o, \hat{M}_1, \ldots, \hat{M}_{K-1}\}$, where $\hat{M}_k = n(\hat{\sigma}_k^2 - \hat{\sigma}^2_{k+1})/\hat{\sigma}^2_{k+1}$, $k = 0, 1, 2, \ldots, K-1$, $\hat{\sigma}_k^2$ is the residual variance corresponding to the model with parameter estimates $\textcircled{H}_k$, and $\hat{\sigma}^2_{k+1}$ is the residual variance corresponding to the model with parameter estimates $\hat{\textcircled{H}}_{k+1}$.

2. For a given choice of α, determine C_k such that $P\{M_k > C_k\} = \alpha$, where M_k is the maximum order statistic in a sequence of (K-k) independent χ^2_1 random variables.

3. Choose the estimated order, $\hat{p}$, such that

$$\hat{p} = \{\min k{:}\hat{M}_k \leq C_k,\ 0 \leq k \leq K\} ,$$

and thus $\hat{\textcircled{H}} = \hat{\textcircled{H}}_{\hat{p}}$. This method of estimating $\textcircled{H}$ is referred to as the "max-χ^2 technique."

To illustrate the application of the max-χ^2 technique in tests for serial correlation, consider the first order AR model with nonzero but unknown lag j_1:

$$Z_t = \emptyset_{j_1} Z_{t-j_1} + \varepsilon_t \quad ,$$

where the $\{\varepsilon_t\}$ is white noise. The hypotheses of interest to the researcher who is attempting to detect serial correlation in $\{Z_t\}$ are then

$$H_o: \emptyset_{j_1} = 0$$
$$H_1: \emptyset_{j_1} \neq 0.$$

Rewriting H_o and H_1 in terms of the parameter set $\textcircled{H}$, we have

$$H_o: \textcircled{H} = \{0\} \text{ (white noise)}$$
$$H_1: \textcircled{H} = \{1;\ j_1\} \ ,$$

where $1 \leq j_1 \leq K$. The test statistic takes the form:

$$\hat{T}_o = n\left(\frac{\hat{\sigma}_o^2 - \hat{\sigma}_1^2}{\hat{\sigma}_1^2}\right)$$

where

$\hat{\sigma}_o^2 = \frac{1}{n}\sum_{t=1}^{n} Z_t = C_o$, and σ_1^2 is generic for the estimated residual variance of each first order AR model. However, McClave has shown that the subset algorithm, when applied to (3.2.2) will choose that lag j_1, which minimizes σ_1^2 over all first order models. Hence, the test statistic takes the form:

$$\hat{M}_o = \max_{j_1} \hat{T}_o = n\left(\frac{\hat{\sigma}_o^2 - \min\limits_{j_1} \hat{\sigma}_1^2}{\min\limits_{j_1} \hat{\sigma}_1^2}\right)$$

Now $\hat{\sigma}_1^2 = C_o(1-r_{j_1}^2)$, where r_{j_1} is the j_1^{th} sample autocorrelation function. Then,

$$\hat{M}_o = n \left\{ \frac{C_o - \min_{j_1}[C_o(1-r_{j_1}^2)]}{\min_{j_1}[C_o(1-r_{j_1}^2)]} \right\} \tag{3.2.3}$$

$$= n \left\{ \frac{C_o - C_o[1-\max_{j_1} r_{j_1}^2]}{C_o[1-\max_{j_1} r_{j_1}^2]} \right\}$$

$$= n \cdot C_o \left\{ \frac{\max_{j_1} r_{j_1}^2}{1-\max_{j_1} r_{j_1}^2} \right\} .$$

In equation (3.2.3), notice that M_o is equivalent to the statistic $n \cdot \max_{j_1} r_{j_1}^2$. McClave showed that $\hat{T}_o \xrightarrow[Ho]{D} \chi_1^2$, i.e., the statistic $\hat{T}_o$ converges in distribution when H_o is true to a χ^2 random variable with 1 degree of freedom. It follows then, that under the hypothesis of white noise, $\hat{M}_o \xrightarrow[H_o]{D} M$, where M is distributed as maximum order statistic from a sequence of K independent χ_1^2 random variables. Values of the test statistic "too large" will lead to a rejection of H_o, where "too large" is determined according to the distribution of M.

Let us compare the "max-χ^2" statistic,

$$n \max_{1 \le j \le K} r_j^2 ,$$

with the Box-Pierce "portmanteau" statistic,

$$n \sum_{j=1}^{K} r_j^2 .$$

Both statistics are based on the vector of sample autocorrelations, $\underset{\sim}{r} = (r_1, r_2, \ldots, r_K)'$. Models (3.1.5) and (3.2.1) both take the form of the more general alternative error structure

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_m Z_{t-m} + \varepsilon_t, \qquad (3.2.4)$$

which gives rise to the plausibility that both testing procedures attain high power against non-first order alternatives. The similarities between the two statistics are evident, and considering the fact that the well-known B-P procedure has been in the literature and widely used since 1970, some may question the motivation of those investigating the max-X^2 technique as a test for serial correlation. Obtaining an answer to this query is the goal of this research. As a partial answer, we offer the following: the max-X^2 test is presented as an alternative testing procedure to those researchers concerned with studying parsimonious models, i.e., models with few nonzero lag parameters. In most practical cases, if the $\{Z_t\}$ are in fact correlated, a minimum number of the lag parameters $(\emptyset_1, \emptyset_2, \ldots, \emptyset_m)$ in model (3.2.4) are nonzero. Consider the researcher who utilizes

$$n \sum_{j=1}^{K} r_j^2$$

as a test statistic. We will discover in Section 3.3 that knowledge of the sample autocorrelation vector $\underset{\sim}{r}$ is sufficient for calculation of the parameter estimates $(\hat{\emptyset}_1, \hat{\emptyset}_2, \ldots, \hat{\emptyset}_m)$. Then the researcher is essentially computing estimates of lag parameters which are known to be zero as well as estimates of those nonzero lag parameters. These "zero" estimates, of course, contribute very little to the statistic

$$n \sum_{j=1}^{K} r_j^2 \quad .$$

Our conjecture is that a predominance of small r_j's will deflate the sum, resulting in a failure to reject the null hypothesis of white noise when in fact several of the true autocorrelations are nonzero. This, in turn, leads to an invalid implementation of the least squares inferential techniques (see Section 1.1). Our goal is to show that, in many cases, the practical research worker can avoid this problem by employing the max-X^2 procedure.

Consideration is given to both the null and non-null asymptotic distributions of the test statistics in the following section. These results become the foundation for the power comparisons discussed in Chapter 4.

3.3 Asymptotic Distribution of the Test Statistics

3.3.1 Preliminaries

Assume that $Z_1, Z_2, \ldots, Z_n$ are observations generated at n consecutive time points by a stationary AR model of order p with lag coefficients $(\emptyset_1, \emptyset_2, \ldots, \emptyset_p)$, i.e.,

$$Z_t = \emptyset Z_{t-1} + \emptyset_2 Z_{t-2} + \ldots + \emptyset_p Z_{t-p} + \varepsilon_t, \quad t = 1, 2, \ldots, n \ , \qquad (3.3.1)$$

where $\{\varepsilon_t\}$ is an uncorrelated series with mean 0 and variance σ^2(white noise). Let

$$\gamma_\nu = E\{Z_t Z_{t+\nu}\} \quad , \quad \nu = 0, \pm 1, \pm 2, \ldots$$

and

$$C_\nu = \frac{1}{n} \sum_{t=1}^{n-|\nu|} Z_t Z_{t+|\nu|}, \quad \nu = 0, \pm 1, \pm 2, \ldots, \pm(n-1),$$

represent the true and sample autocovariance functions, respectively. Also, let $\rho_\nu = \gamma_o^{-1}\gamma_\nu$ and $r_\nu = C_o^{-1}C_\nu$ be the true and sample autocorrelations of order ν, respectively. Then the vector of lag coefficients $\underset{\sim}{\emptyset} = (\emptyset_1, \emptyset_2, \ldots, \emptyset_p)'$ is estimated by $\hat{\underset{\sim}{\emptyset}} = (\hat{\emptyset}_1, \hat{\emptyset}_2, \ldots, \hat{\emptyset}_p)'$ in the modified

Yule-Walker equations:

$$\hat{\underset{\sim}{\Gamma}}\hat{\underset{\sim}{\emptyset}} = \underset{\sim}{r} \quad ,$$

where $\hat{\Gamma}$ is a symmetric pxp matrix given by:

$$\hat{\Gamma} = \begin{bmatrix} 1 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & 1 & r_1 & \cdots & r_{p-2} \\ r_2 & r_1 & 1 & \cdots & r_{p-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{p-1} & & & & 1 \end{bmatrix} \quad ,$$

and $\underset{\sim}{r} = (r_1, r_2, \ldots, r_p)'$. Also, the estimate of σ^2 is given by

$$\hat{\sigma}_p^{\,2} = C_o[1 - r'\hat{\Gamma}r].$$

3.3.2 General Case

Due to the complexities involved with obtaining the exact distribution of the two test statistics of interest, this research will consider the large sample properties of the tests. Anderson [4], Grenander and Rosenblatt [15], Fuller [13] and others showed that the asymptotic joint distribution of the first K sample autocorrelations is multivariate normal, although Bartlett [6] was one of the first to consider the form of the limiting variance of the sample autocorrelation. This asymptotic result, given in its most general form by Anderson, is stated in the lemma that follows.

Lemma 1: Let

$$Z_t = \sum_{j=-\infty}^{\infty} \alpha_j \varepsilon_{t-j} \quad ,$$

where $\{\varepsilon_t\}$ is white noise,

$$\sum_{j=-\infty}^{\infty} |\alpha_j| < \infty \quad ,$$

and

$$\sum_{j=-\infty}^{\infty} |j| \alpha_j^2 < \infty \quad .$$

Define the quantity P_ν, as

$$P_\nu = \sum_{j=-\infty}^{\infty} \rho_j \rho_{\nu-j} \quad ,$$

where

$$\rho_\nu = \gamma_o^{-1} \gamma_\nu = E\{Z_t Z_{t+\nu}\}/E\{Z_t^2\} \quad ,$$

and let

$$r_\nu = \frac{\sum_{t=\nu+1}^{n} Z_t Z_{t-\nu}}{\sum_{t=1}^{n} Z_t^2} \quad , \quad 1 \leq \nu \leq K \quad .$$

Then $\sqrt{n}(\underset{\sim}{r} - \underset{\sim}{\rho}) \xrightarrow{D} N_K\{\underset{\sim}{o}, V\}$, where $\underset{\sim}{r} = (r_1, r_2, \ldots, r_K)'$, $\underset{\sim}{\rho} = (\rho_1, \rho_2, \ldots, \rho_K)'$ and $V = (v_{ij})$ is the KxK dispersion matrix of the limiting distribution whose elements are given by:

$$v_{ii} = P_o + P_{2i} + 2\rho_i^{\,2} P_o - 4\rho_i P_i \quad , \quad 1 \leq i \leq K \qquad (3.3.2)$$

$$v_{ij} = P_{i+j} + P_{i-j} + 2\rho_i \rho_j P_o - 2\rho_i P_j - 2\rho_j P_i \quad ,$$
$$1 \leq j < i < K \ . \qquad (3.3.3)$$

Proof: See Anderson [4], pp. 489-495. □

Recall the invertibility property of the ARMA model in Section 3.1. Fuller [13] and others show that $\{Z_t\}$ generated from the AR process given by (3.3.1) can be expressed as an infinite moving average (MA) of $\{\varepsilon_t\}$, i.e.,

$$Z_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} \quad ,$$

with $\{\alpha_j\}$ absolutely summable, if the roots of the equation

$$x^p + \emptyset_1 x^{p-1} + \emptyset_2 x^{p-2} + \ . \ . \ . + \emptyset_p = 0$$

are all in the unit circle. Since model (3.3.1) is stationary, this condition is satisfied and thus the AR process is just a special case of the time series model defined in Lemma 1. However, before applying Lemma 1 to obtain the asymptotic distribution of the Box-Pierce and max-χ^2 test statistic, we point out that in practice, the $\{Z_t\}$ in model (3.3.1) are unknown. Hence, the researcher, unable to calculate the sample autocorrelations,

$$r_\nu = \frac{\sum_{t=\nu+1}^{n} Z_t Z_{t-\nu}}{\sum_{t=1}^{n} Z_t^2} \qquad 1 \leq \nu \leq K \quad ,$$

must, instead, calculate the estimated sample autocorrelations,

$$\hat{r}_\nu = \frac{\sum_{t=\nu+1}^{n} \hat{Z}_t \hat{Z}_{t-\nu}}{\sum_{t=1}^{n} \hat{Z}_t^2} \quad , \qquad 1 \leq \nu \leq K \quad ,$$

Where $\{\hat{Z}_t\}$ are the residuals computed from least squares regression. The $\{\hat{r}_\nu\}$ are used in place of $\{\hat{r}_\nu\}$ when calculating the Box-Pierce and max-χ^2 test statistics. Thus, our interest lies in determining the limiting behavior of the estimated sample autocorrelations, $(\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_k)$.

A result due to Anderson [4], formally stated in Lemma 2, shows that under certain regularity conditions on the design matrix X of regression, the limiting distribution of the sample autocorrelations defined in Lemma 1 also holds for autocorrelations computed using the residuals $\{\hat{Z}_t\}$ calculated from least squares regression.

Lemma 2: Let $\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{Z}$ be the regression model given in equation (1.1.1), where

$$Z_t = \sum_{j=-\infty}^{\infty} \alpha_j \, \varepsilon_{t-j} \, ,$$

$$\sum_{j=-\infty}^{\infty} |\alpha_j| < \infty \quad ,$$

$$\sum_{j=-\infty}^{\infty} |j| \alpha^2 < \infty \quad ,$$

and where $\{\varepsilon_t\}$ is white noise with bounded moments. Let

$$a_{ij}^n(h) = \sum_{t=1}^{n-h} x_{i,t+h} \, x_{jt}, \quad 1 \le i \le g, \; 1 \le j \le g$$

where

$$X = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{g1} \\ x_{12} & x_{22} & \cdots & x_{g2} \\ \vdots & & & \\ x_{1n} & x_{2n} & \cdots & x_{gn} \end{bmatrix}$$

Given that the following four conditions (Grenander's condition) hold:

(1) $a_{ii}^{n}(0) \to \infty$ as $n \to \infty$

(2) $\dfrac{x_{i,n+1}^{2}}{a_{ii}^{n}(0)} \to 0$ as $n \to \infty$

(3) $\dfrac{a_{ij}^{n}(h)}{\sqrt{a_{ii}^{n}(0)a_{jj}^{n}(0)}} \to \rho_{ij}^{*}(h)$ as $n \to \infty$

(4) $R(0)$ is nonsingular, where $R(h) = [\rho_{ij}^{*}(h)]$,

then

$$\sqrt{n}(\hat{\underset{\sim}{r}} - \underset{\sim}{\rho}) \xrightarrow{D} N_K \ \{\underset{\sim}{0},\ V\} ,$$

where $\underset{\sim}{\rho}$ and V are defined in Lemma 1,

$$\hat{r}_j = \frac{\sum_{t=j+1}^{n} Z_t Z_{t-j}}{\sum_{t=1}^{n} \hat{Z}_t^2} \quad , \qquad 1 \leq j \leq K,$$

and $\hat{\underset{\sim}{Z}} = \underset{\sim}{Y} - X\hat{\underset{\sim}{\beta}} = [I-X(X'X)^{-1}X^{1}]\underset{\sim}{Y}$.

Proof: See Anderson [4] , pg. 593. □

In order to simplify notation, let

$$T^{(BP)} = n\sum_{j=1}^{K} r_j^2 = n\underset{\sim}{r}'\underset{\sim}{r}$$

and

$$T^{(m)} = n \max_{1 \leq j \leq K} r_j^2 \quad .$$

The asymptotic null and non-null distributions of $T^{(BP)}$ and $T^{(m)}$ are given in theorem form. As a consequence of Lemma 2, the results stated throughout the remainder of this chapter may also be applied with $T^{(BP)}$ and $T^{(m)}$ are computed from least square residuals.

Theorem 1: Let $\underset{\sim}{r} = (r_1, r_2, \ldots, r_K)'$ be the vector of sample autocorrelations generated from model (3.3.1), and let $T^{(BP)}$ and $T^{(m)}$

represent the Box-Pierce and max-χ^2 test statistics, respectively.

Then

$$T^{(BP)} \xrightarrow[H_o]{D} \chi^2_K(o),$$

i.e., $T^{(BP)}$ converges in distribution when the null hypothesis of white noise is true to a central χ^2 random variable with K degrees of freedom, and

$$T^{(m)} \xrightarrow[H_o]{D} M,$$

i.e., $T^{(m)}$ converges in distribution when the null hypothesis of white noise is true to the distribution of M, where M represents the maximum order statistic from a sequence of K independent χ^2_1 random variables.

Proof: Under the null hypothesis of white noise, $\{Z_t\}$ is uncorrelated with mean 0 and variance σ^2. Without loss of generality, take $\sigma^2 = 1$. In the null case, we have then

$$\gamma_\nu = \begin{cases} 1 & \nu = 0 \\ 0 & \nu = 1, 2, \ldots, K \end{cases},$$

$$\rho_\nu = \begin{cases} 1 & \nu = 0 \\ 0 & \nu = 1, 2, \ldots, K \end{cases},$$

and

$$P_\nu = \begin{cases} 1 & \nu = 0 \\ 0 & \nu = 1, 2, \ldots, K \end{cases}.$$

Recall that $\{Z_t\}$ in the stationary AR model of (3.3.1) may be written as an infinite MA process,

$$Z_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}$$

where the $\{\alpha_j\}$ are absolutely summable. Having satisfied the conditions of Lemma 1, we apply equations (3.3.2) and (3.3.3) to obtain

$$v_{ij} = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases} \quad (i=1,2,\ldots,K),(j=1,2,\ldots,K) .$$

Hence, by Lemma 1:

$$\sqrt{n}\,\underset{\sim}{r} \xrightarrow[H_o]{D} N_K\{\underset{\sim}{0},I\} .$$

Then $T^{(BP)} \xrightarrow[H_o]{D} \chi^2_K(0)$, where $\chi^2_K(o)$ is a central χ^2 random variable with K degrees of freedom, and $T^{(m)} \xrightarrow[H_o]{D} M$, where M is the maximum order statistic from a sequence of K independent χ^2_1 random variables. ☐

We now proceed to find the appropriate critical values, $C_\alpha^{(BP)}$ and $C_\alpha^{(m)}$, where:

$$P_{H_o}\{T^{(BP)} > C_\alpha^{(BP)}\} = \alpha ,$$

and

$$P_{H_o}\{T^{(m)} > C_\alpha^{(m)}\} = \alpha .$$

It is obvious that $C_\alpha^{(BP)} = \chi^2_{K,\alpha}$, where $\chi^2_{K,\alpha}$ is given by

$$P\{\chi^2_K(0) > \chi^2_{K,\alpha}\} = \alpha .$$

In order to find $C_\alpha^{(m)}$, let $C_\alpha^{(m)} = c$ and observe that:

$$P\{n \max_{1\leq j\leq K} r_j^2 < c\} = P\{nr_1^2 < c, nr_2^2 < c, \ldots, nr_K^2 < c\}$$

$$= P\{nr_1^2 < c\}\cdot P\{nr_2^2 < c\}\cdots P\{nr_K^2 < c\}$$

$$= P\{-\sqrt{c} < \sqrt{n}r_1 < \sqrt{c}\}\cdot P\{-\sqrt{c} < \sqrt{n}r_2 < \sqrt{c}\}\cdots P\{-\sqrt{c} < \sqrt{n}r_K < \sqrt{c}\}$$

$$= [P\{-\sqrt{c} < Z < \sqrt{c}\}]^K ,$$

where Z is a normal random variable with mean 0 and variance 1. Then

$$1-\alpha = [P\{-\sqrt{c} < Z < \sqrt{c}\}]^K$$

implies

$$(1-\alpha)^{\frac{1}{K}} = P\{-\sqrt{c} < Z < \sqrt{c}\}$$

$$= 2\Phi(\sqrt{c})-1,$$

where

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} d_z.$$

Thus, we have that

$$c_\alpha^{(m)} = [\Phi^{-1}\{\frac{1}{2} + \frac{(1-\alpha)^{1/K}}{2}\}]^2$$

The two procedures for testing $H_o:\{Z_t\}$ uncorrelated versus $H_a:\{Z_t\}$ correlated, at significance level α, can be outlined as follows:

1. Box-Pierce test: reject H_o in favor of H_1 if $T^{(BP)} > \chi^2_{K,\alpha}$;
2. Max-χ^2 test: reject H_o in favor of H_1 if

$$T^{(m)} > [\Phi^{-1}\{\frac{1}{2} + \frac{(1-\alpha)^{1/K}}{2}\}]^2 .$$

In order to compare the two procedures in terms of power (Chapter 4), consideration must also be given to the non-null distributions of the test statistics. However, the asymptotic distributions of $T^{(BP)}$ and $T^{(m)}$ are not so easily derived when the $\{Z_t\}$ are dependent.

Recall that $\sqrt{n}(\underset{\sim}{r} - \underset{\sim}{\rho}) \xrightarrow{D} N_K\{\underset{\sim}{0},V\}$ (Lemma 1). When the alternative of serial correlation is true, the asymptotic dispersion matrix V has a non-diagonal form. For our purposes, we write that, under H_a, $\sqrt{n}\ \underset{\sim}{r} \overset{\cdot}{\underset{\cdot}{\sim}} N_K\{\sqrt{n}\ \underset{\sim}{\rho},V\}$ for large n, where $\overset{\cdot}{\underset{\cdot}{\sim}}$ means "has an approximate distribution," and $\underset{\sim}{\rho}$ and V are determined from the expressions given in Lemma 1. A result due to H. Scheffé[28], given here as a lemma, enables us to express the Box-Pierce statistic $T^{(BP)}$ as a linear combination of independent non-central χ^2 random variables. Following this lemma, the asymptotic non-null distribution of $T^{(BP)}$ is given in theorem form.

Lemma 3: Let $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$ be a column random vector which follows a multivariate normal law with mean vector $\underset{\sim}{\mu}$ and covariance matrix Σ. Consider the quadratic form $Q = x'\Lambda x$, where Λ is a KxK symmetrix matrix. Then if Σ is nonsingular, Q may be expressed in the form:

$$Q = \sum_{i=1}^{\ell} \lambda_i \chi^2_{r_i}(\delta_i^2),$$

where $(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ are the distinct nonzero eigenvalues of $\Lambda\Sigma$, $(r_1, r_2, \ldots, r_\ell)$ their respective orders of multiplicity, $(\delta_1^2, \delta_2^2, \ldots, \delta_\ell^2)$ are certain linear combinations of $(\mu_1, \mu_2, \ldots, \mu_n)$, and the $\{\chi^2_{r_i}(\delta_i)\}_{i=1}^{\ell}$ a sequence of independent non-central χ^2 random variables with r_i degrees of freedom and non-centrality parameter δ_i^2.

Proof: See Scheffé [28], pp. 418. □

Theorem 2: Let $\underset{\sim}{r} = (r_1, r_2, \ldots, r_K)'$ be the vector of sample autocorrelations generated from model (3.3.1), and $V = (v_{ij})$ the symmetric positive definite KxK dispersion matrix of the asymptotic distribution of $\sqrt{n}\,\underset{\sim}{r}$. Also, let $V = TT'$, T a lower triangular matrix, and consider the statistic

$$T^{(BP)} = n\sum_{j=1}^{K} r_j^2 = n(\underset{\sim}{r}'\underset{\sim}{r}).$$

Then we can write, for large n:

$$T^{(BP)} = \sum_{i=1}^{\ell} \lambda_i \chi^2_{r_i}(\delta_i^2),$$

where (i) $(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ are the distinct nonzero eigenvalues of V,

(ii) $(m_1, m_2, \ldots, m_\ell)$ are the respective orders of multiplicities for the eigenvalues,

(iii) $\{\chi^2_{m_i}(\delta_i^2)\}_{i=1}^{\ell}$ a sequence of independent non-central χ^2 random variables with m_i degrees of freedom and non-centrality parameter δ_i^2, $i = 1, 2, \ldots, \ell$,

(iv) $\delta_i^2 = n\underset{\sim}{\rho}'V^{-1}T\cdot\{\prod_{\substack{s=1\\ s\neq i}}^{\ell}(\frac{1}{\lambda_i-\lambda_s})\cdot(T'T-\lambda_s\cdot I)\}\cdot T^{-1}\underset{\sim}{\rho}$, $i = 1, 2, \ldots, \ell$.

Proof: From Lemma 1, we have $\sqrt{n}\,\underset{\sim}{r} \overset{\cdot}{\sim} N_K\{\sqrt{n}\,\underset{\sim}{\rho}, V\}$ for large n. Let $\underset{\sim}{x} = \sqrt{n}\,\underset{\sim}{r}$, $\underset{\sim}{\mu} = \sqrt{n}\,\underset{\sim}{\rho}$, $\Sigma = V$, and $\Lambda = I$ in Lemma 3. Then

$$T^{(BP)} = n(\underset{\sim}{r}'\underset{\sim}{r}) = (\sqrt{n}\,\underset{\sim}{r})'I\,(\sqrt{n}\,\underset{\sim}{r}) = \underset{\sim}{x}'\Lambda\,\underset{\sim}{x} \quad ,$$

and the conditions of Lemma 3 are satisfied. Thus, for large n,

$$T^{(BP)} = \sum_{i=1} \lambda_i \chi^2_{r_i}(\delta_i^2) \quad ,$$

where λ_i, m_i, and $\chi^2_{m_i}(\delta_i^2)$ are defined by (i), (ii), and (iii). If V is positive definite, W.Y. Tan [31] obtained a formula useful for computing the non-centrality parameters $\{\delta_i^2\}$:

$$\delta_i^2 = n\,\underset{\sim}{\rho}'\cdot V^{-1}\cdot T\cdot E_i\cdot T^{-1}\cdot\underset{\sim}{\rho} \ , \qquad (i = 1, 2, \ldots, \ell),$$

where $V = TT'$ and $E_i = \prod_{\substack{s=i\\ s\neq i}}^{\ell} (\lambda_i-\lambda_s)^{-1}\cdot(T'T-\lambda_s\cdot I)$, $i = 1, 2, \ldots, \ell$.

This is condition (iv) of theorem, and thus completes the proof. □

By Theorem 2, the asymptotic distribution of $T^{(BP)}$ is found by determining the distribution of a weighted sum of independent non-central χ^2 random variables. The problem is not an easy one, but J. P. Imhoff [23] and others have succeeded in deriving an expression for the distribution of the quadratic form Q (given in Lemma 4), by means of an inverstion of the characteristic function of Q. These results are given in Section 3.3.3.2.1 and later applied in Chapter 4, where we seek powers of the test.

Similarly, the asymptotic non-null distribution of the max-χ^2 statistic, $T^{(m)}$ is a complex one as described in the following theorem.

Theorem 3: Let $\underset{\sim}{r} = (r_1, r_2, \ldots, r_K)'$ be the vector of the sample autocorrelations generated from model (3.3.1), and let $T^{(m)} = n \cdot \max_{1\leq j\leq K} r_j^2$. Then, for large n, $T^{(m)}$ is approximately distributed as the maximum order statistic from a sequence of K dependent non-central χ_1^2 random variables, $\{v_{ii} \cdot \chi_1^2(\delta_i^2)\}_{i=1}^{\ell}$ with

(i) non-centrality parameters given by $\delta_i^2 = \dfrac{n\rho_i^2}{2v_{ii}}$, $1\leq i\leq K$

and

(ii) $v_{ii} = P_o + P_{2i} + 2\rho_i^2 P_0 - 4\rho_i P_i$, $1\leq i\leq K$.

Proof: Again, by Lemma 1, $\sqrt{n}\,\underset{\sim}{r} \mathrel{\dot{\sim}} N_K\{\sqrt{n}\,\underset{\sim}{\rho}, V\}$ for large n, where the elements of $V=(v_{ij})$ are found from equations (3.3.2) and (3.3.3). Hence, for large n, the marginal distributions of $\{\sqrt{n}\, r_i\}_{i=1}^{K}$ can be approximated by a sequence of K dependent normal random variables with means $\{\sqrt{n}\rho_i\}_{i=1}^{K}$ and variances $\{v_{ii}\}_{i=1}^{K}$, respectively. Now, for $1\leq i\leq K$, $\sqrt{n}\, r_i \sim N\{\sqrt{n}\,\rho_i, v_{ii}\}$ implies that

$$\frac{\sqrt{n}\, r_i}{\sqrt{v_{ii}}} \sim N\left\{\frac{\sqrt{n}\,\rho_i}{\sqrt{v_{ii}}}, 1\right\} .$$

Thus, we have that $\dfrac{n r_i^2}{v_{ii}} \sim \chi_1^2(\delta_i^2)$, where $\delta_i^2 = \dfrac{n\rho_i^2}{2v_{ii}}$; or, we write, $n r_i^2 \sim v_{ii} \cdot \chi_1^2(\delta_i^2)$, $1\leq i\leq K$. It follows that $T^{(m)} = \max_{1\leq j\leq K}(n r_j^2)$ is approximately distributed as the maximum order statistic of the dependent (since $v_{ij} \neq 0$ in general) sequence $\{v_{ii}\chi_1^2(\delta_i^2)\}_{i=1}^{K}$. □

The complexity of the distribution derived in Theorem 3 makes power computations very difficult. In Section 3.3.3.2.2, this problem is made somewhat more workable by considering a Taylor-series expansion of the density function of $T^{(m)}$.

In the sections that follow, a few simple alternative error models are presented, and the non-null distributions of the max-χ^2 and Box-Pierce statistics are discussed. The severity of the distributional problem, even in these simple cases, will be illustrated.

3.3.3 Special Cases

Recall the first order AR alternative error model considered by Durbin-Watson (Section 1.2.1). Let us consider a more general form of this model,

$$Z_t = \emptyset Z_{t-J} + \varepsilon_t \ , \qquad (3.3.4)$$

where t = 1, 2, . . ., n and $\{\varepsilon_t\}$ a white noise process. In the terminology of the previous sections, we call model (3.3.4) a first order AR model with maximum (and singleton) lag J, where J is unknown.

Lemma 4: The vector of true autocorrelations, $\underset{\sim}{\rho} = (\underset{\sim}{\rho}_1, \underset{\sim}{\rho}_2, \ldots)'$, for model (3.3.4) are calculated as follows:

$$\rho_\nu = \begin{cases} 1 \ \ldots & \nu = 0 \\ \emptyset^m \ \ldots & \nu = mJ;\ m = 1, 2, 3, \ldots \\ 0 \ \ldots & \text{otherwise} \end{cases} \qquad (3.3.5)$$

Proof: Recall that $\rho_\nu = \gamma_\nu/\gamma_o$, $\nu = 1, 2, 3, \ldots$. Using model (3.3.4) we write for $\nu = 1, 2, \ldots$

$$\gamma_\nu = E\{Z_t \cdot Z_{t-\nu}\} = E\{\phi Z_{t-\nu} Z_{t-J} + \varepsilon_t Z_{t-\nu}\}$$

$$= \phi E\{Z_t Z_{t-(J-\nu)}\} + E\{\varepsilon_t Z_{t-\nu}\}.$$

Now, $E\{\varepsilon_t Z_{t-\nu}\} = 0$ since $Z_{t-\nu}$ is uncorrelated with $\varepsilon_{t-\nu+1}$, $\varepsilon_{t-\nu+2}$, . . ., ε_t, ε_{t+1}, ... Hence,

$$\gamma_\nu = \phi \cdot E\{Z_t \cdot Z_{t-(J-\nu)}\} = \phi \cdot \gamma_{J-\nu}\ , \quad (\nu = 1, 2, 3, \ldots). \quad (3.3.6)$$

Immediately, we see that for m = 1, 2, 3, . . .,

$$\gamma_{mJ} = \phi \gamma_{(m-1)J}\ ,$$

and thus,

$$\rho_{mJ} = \phi \rho_{(m-1)J}\ .$$

Since $\rho_o = 1$, it follows that $\rho_{mJ} = \phi^m$, (m = 1, 2, 3, . . .). Consider now, those autocorrelations of order not a multiple of J. Take $\nu = J-\nu$ in equation (3.3.6). Then $\gamma_{J-\nu} = \phi\gamma_\nu$; but also $\gamma_\nu = \phi\gamma_{J-\nu}$. This implies that $\gamma_\nu = \phi^2\gamma_\nu$, for ν = 1, 2, 3, . . ., J-1, J+1, . . ., 2J-1, 2J+1, ... And since $\phi \neq 0$, it must be that $\gamma_\nu = 0$ for $\nu \neq mJ$, (m = 1, 2, 3, . . .). This completes the proof. □

The vector $\underset{\sim}{\rho}$, given in equation (3.3.5), will be used to compute the asymptotic mean vector of the sample autocorrelations for this special case.

Lemma 5: For the model (3.3.4), the quantities $\{P_\nu\}$ ν = 1, 2, . . ., defined in Lemma 1, are given by:

$$P_\nu = \begin{cases} \dfrac{1+\phi^2}{1-\phi^2} & \ldots \quad \nu = 0 \\[2ex] \dfrac{\phi^m[m+1-(m-1)\phi^2]}{1-\phi^2} & \nu = mJ, \ (m = 1, 2, 3, \ldots) \\[2ex] 0 & \text{otherwise} \end{cases} \quad (3.3.7)$$

Proof: From Lemma 1, we have $P_\nu = \sum_{j=-\infty}^{\infty} \rho_j \rho_{\nu-j}$, $(\nu = 0, \pm 1, \pm 2, \ldots)$. First let $\nu = 0$. Then we have

$$P_o = \sum_{j=-\infty}^{\infty} \rho_j^2 = \rho_o + 2\sum_{j=1}^{\infty} \rho_j^2 \quad .$$

Applying the results from Theorem 4, we obtain

$$P_o = 1 + 2\sum_{m=1}^{\infty} \rho_{mJ}^2 = 1 + 2\sum_{m=1}^{\infty} \phi^{2m} \; .$$

Since

$$\sum_{i=1}^{\infty} a^i = \frac{a}{1-a}$$

if $|a| < 1$, we have

$$P_o = 1 + 2\left(\frac{\phi^2}{1-\phi^2}\right) = \frac{1+\phi^2}{1-\phi^2} \quad .$$

For ν a non-multiple of J and $j \varepsilon (-\infty, \infty)$, the integers j and $\nu - j$ cannot both be multiples of J simultaneously. It follows that $P_\nu = 0$ for $\nu \neq mJ$, $(m = 1, 2, \ldots)$. To find the quantities P_{mJ}, $(m = 1, 2, \ldots)$, we write

$$\begin{aligned} P_{mJ} &= \sum_{j=-\infty}^{\infty} \rho_j \rho_{mJ-j} = \sum_{j=-\infty}^{\infty} \rho_{jJ} \rho_{(m-j)J} \\ &= \sum_{j=-\infty}^{-1} \rho_{jJ} \rho_{(m-j)J} + \sum_{j=0}^{m} \rho_{jJ}\ {}_{(m-j)J} + \sum_{j=m+1}^{\infty} \rho_{jJ} \rho_{(m-j)J} . \end{aligned}$$

Utilizing the results in Theorem 4, we have

$$P_{mJ} = \sum_{j=1}^{\infty} \phi^j \phi^{m+j} + \sum_{j=0}^{m} \phi^j \phi^{m-j} + \sum_{j=m+1}^{\infty} \phi^j \phi^{j-m} \; . \qquad (3.3.8)$$

Now take $i=j-m$ in the last summation of equation (3.3.8) and simplify to finally obtain:

$$P_{mJ} = \phi^m \sum_{j=1}^{\infty} \phi^{2j} + (m+1)\phi^m + \phi^m \sum_{i=1}^{\infty} \phi^{2i}$$

$$= 2\phi^m \sum_{j=1}^{\infty} \phi^{2j} + (m+1)\phi^m$$

$$= 2\phi^m \left(\frac{\phi^2}{1-\phi^2}\right) + (m+1)\phi^m$$

$$= \phi^m[m+1 - (m-1)\phi^2]/(1-\phi^2). \quad \square$$

We now proceed to find the structure of the asymptotic distribution of the sample autocorrelations for this special case using the results derived in Lemmas 4 and 5.

Theorem 4: The covariance structure for the asymptotic distribution of the first K sample autocorrelations, $(r_1, r_2, \ldots, r_K)$, from model (3.3.4), is given by:

$$v_{\nu\nu} = \begin{cases} \dfrac{1+\phi^2-(2m+1)\phi^{2m}+(2m-1)\phi^{2(m+1)}}{1-\phi^2} & \ldots \quad \nu = mJ \; (m = 1, 2, 3, \ldots) \\[2ex] \dfrac{1+\phi^2+(m+1)\phi^m-(m-1)\phi^{m+2}}{1-\phi^2} & \ldots \quad \nu = \dfrac{mJ}{2}, J \text{ even } (m = 1, 3, 5, \ldots) \\[2ex] \dfrac{1+\phi^2}{1-\phi^2} & \ldots \quad \text{otherwise} \end{cases} \tag{3.3.9}$$

$$v_{\nu_1,\nu_2} = \begin{cases} \dfrac{\phi^m[m+1-(m-1)\phi^2]}{1-\phi^2} & \ldots \quad \nu_1+\nu_2=mJ \text{ or } \nu_1-\nu_2 = mJ \text{ but not both } (m = 1, 2, 3, \ldots) \\[2ex] \dfrac{\phi^m[m+1-(m-1)\phi^2]+\phi^\ell[\ell+1-(\ell-1)\phi^2]}{1-\phi} & \ldots \quad \nu_1+\nu_2=mJ \; (m=1,2,3\ldots), \; \nu_1-\nu_2=\ell J \; (\ell=1,2,3\ldots) \\[2ex] \dfrac{\phi^{m-\ell}[m-\ell+1(m-\ell-1)\phi^2]-\phi^{m+\ell}[m+\ell+1-(m+\ell-1)\phi^2]}{1-\phi^2} & \ldots \quad \nu_1 = mJ \; (m = 1, 2, \ldots), \; \nu_2 = \ell J \; (\ell = 1, 2, \ldots) \\[2ex] 0 & \quad \text{otherwise} \end{cases} \tag{3.3.10}$$

Proof: Recall from Lemma 1, the expression for the asymptotic variance of the ν^{th} order sample autocorrelation, $\sqrt{n}\, r_\nu$:

$$v_{\nu\nu} = P_o + P_{2\nu} + 2\rho_\nu^2 P_o - 4\rho_\nu P_\nu \;. \qquad (3.3.11)$$

For $\nu \neq mJ$, we have $\rho_\nu = 0$ (Lemma 4) and $P_\nu = P_{2\nu} = 0$ (Lemma 5), thus

$$v_{\nu\nu} = P_o = \frac{1+\phi^2}{1-\phi^2} \qquad (\nu \neq mJ;\ m = 1, 2, 3, \ldots) \;.$$

Now consider $\nu = mJ$. Here we have from Lemmas 5 and 6, $\rho_\nu = \phi^m$,

$$P_\nu = \frac{\phi^m[m+1-(m-1)\phi^2]}{1-\phi^2},$$

and

$$P_{2\nu} = \frac{\phi^{2m}[2m+1-(2m-1)\phi^2]}{1-\phi^2} \;.$$

We then can write

$$\begin{aligned} v_{\nu\nu} &= \frac{1+\phi^2}{1-\phi^2} + \frac{\phi^{2m}[2m+1-(2m-1)\phi^2]}{1-\phi^2} + 2\phi^{2m}\left(\frac{1+\phi^2}{1-\phi^2}\right) - 4\frac{\phi^{2m}[m+1-(m-1)\phi^2]}{1-\phi^2} \\ &= \{1+\phi^2+\phi^{2m}[-2m-1+\phi^2(-1+2m)]\}\ /(1-\phi^2) \\ &= \{1+\phi^2-(2m+1)\phi^{2m}+(2m-1)\phi^{2(m+1)}\}/(1-\phi^2),\ (\nu = mJ;\ m = 1, 2, \ldots). \end{aligned}$$

Because of the term $P_{2\nu}$ in equation (3.3.11), we must also consider those orders of multiples of J/2, when J is an even integer, i.e., $\nu = J/2, 3J/2, 5J/2, \ldots$. Here 2ν becomes an odd multiple of J, thus (from Lemmas 4 and 5)

$$\rho_\nu = 0, P_\nu = 0 \text{ and } P_{2\nu} = P_{mJ} = \frac{\phi^m[m+1-(m-1)\phi^2]}{1-\phi^2},\ (m = 1, 3, 5, \ldots).$$

It follows that

$$v_{\nu\nu} = \frac{1+\phi^2}{1-\phi^2} + \frac{\phi^m[m+1-(m-1)\phi^2]}{1-\phi^2}$$

$$= \frac{1+\phi^2 + (m+1)\phi^m-(m-1)\phi^{m+2}}{1-\phi^2}\ , \quad \nu = \frac{mJ}{2}\ , \text{ J even } (m = 1, 3, 5, \ldots).$$

This completes the proof of equation (3.3.9).

The asymptotic covariances of the sample autocorrelations are found in a similar manner. Rewriting equation (3.3.3) from Lemma 1, we have, for large n,

$$v_{\nu_1,\nu_2} = P_{\nu_1+\nu_2} + P_{\nu_1-\nu_2} + 2\rho_{\nu_1}\rho_{\nu_2}P_o - 2\rho_{\nu_1}P_{\nu_2}$$

$$- 2\rho_{\nu_2}P_{\nu_1}, \ (\nu_1 \neq \nu_2). \qquad (3.3.12)$$

These computations become quite involved, for now we must consider those values of ν_1, ν_2, $\nu_1 + \nu_2$ or $\nu_1 - \nu_2$ which are multiples of J. The result is obtained in a straightforward but nontrivial fashion if we group the possibilities as follows:

a) ν_1 a multiple of J, but ν_2 a nonmultiple of J.

b) ν_2 a multiple of J, but ν_1 a nonmultiple of J.

c) $\nu_1 + \nu_2$ a multiple of J, but $\nu_1 - \nu_2$ a nonmultiple of J.

d) $\nu_1 - \nu_2$ a multiple of J, but $\nu_1 + \nu_2$ a nonmultiple of J.

e) Both $\nu_1 + \nu_2$ and $\nu_1 - \nu_2$ multiples of J, but neither ν_1 nor ν_2 a multiple of J.

f) Both ν_1 and ν_2 multiples of J.

Notice that in case (a) and case (b), neither $\nu_1 + \nu_2$ nor $\nu_1 - \nu_2$ can be a multiple of J. Hence, each of the quantities summed in equation (3.3.12) are zero. Thus,

$$v_{\nu_1,\nu_2} = 0 \text{ for } \nu_1 \neq mJ \text{ or } \nu_2 \neq mJ \quad (m = 1, 2, 3, \ldots)$$

For case (c), the only nonzero term in (3.3.12) is $P_{\nu_1+\nu_2}$. Similarly, for case (d), the only nonzero term in (3.3.12) is $P_{\nu_1-\nu_2}$. This gives

$$v_{\nu_1,\nu_2} = P_{mJ} = \frac{\phi^m[m+1-(m-1)\phi^2]}{1-\phi^2}, \quad \text{for either } \nu_1+\nu_2=mJ$$

or $\nu_1-\nu_2=mJ$ but not both (m = 1, 2, . . .).

When the orders of the sample autocorrelations follow the pattern in case (e), only the terms $P_{\nu_1+\nu_2}$ and $P_{\nu_1-\nu_2}$ are nonzero. Hence,

$$v_{\nu_1,\nu_2} = P_{mJ}+P_{\nu J} = \frac{\phi^m[m+1-(m-1)\phi^2] + \phi^\ell[\ell+1-(\ell-1)\phi^2]}{1-\phi^2},$$

for $\nu_1+\nu_2=mJ$ (m = 1, 2, 3, . . .), and $\nu_1-\nu_2=\ell J$ (ℓ= 1, 2, 3, . . .).

Finally, we consider the order pattern in (f). Both ν_1 and ν_2 multiples of J imply that both $\nu_1+\nu_2$ and $\nu_1-\nu_2$ are multiples of J. Thus, none of the terms summed in (3.3.12) are nonzero. Utilizing Theorems 4 and 5, and lettin $\nu_1=mJ$ and $\nu_2=\ell J$, we obtain:

$$\begin{aligned} v_{\nu_1,\nu_2} &= P_{(m+\ell)J}+P_{(m-\ell)J}+2\rho_{mJ}\rho_{\ell J}P_o-2\rho_{mJ}P_{\ell J}-2\rho_{\ell J}P_{mJ} \\ &= \frac{\phi^{m+\ell}[m+\ell+1-(m+\ell-1)\phi^2]}{1-\phi^2} + \frac{\phi^{m-\ell}[m-\ell+1-(m-\ell-1)\phi^2]}{1-\phi^2} \\ &\qquad + 2\phi^{m+\ell}\left(\frac{1+\phi^2}{1-\phi^2}\right) \\ &\qquad - \frac{2\phi^{m+\ell}[\ell+1-(\ell-1)\phi^2]}{1-\phi^2} - \frac{2\phi^{m+\ell}[m+1-(m-1)\phi^2]}{1-\phi^2} \\ &= \frac{\phi^{m-\ell}[m-\ell+1-(m-\ell-1)\phi^2] - \phi^{m+\ell}[m+\ell+1-(m+\ell-1)\phi^2]}{1-\phi^2} \end{aligned}$$

This completes the proof. $\square$

3.3.3.1. Model 1: The Case J=2, K=2

Let us now apply these results to model (3.3.4) when J=2, i.e.,

$$Z_t = \phi Z_{t-2} + \varepsilon_t \ . \tag{3.3.13}$$

Suppose also that the researcher chooses to base the test of

$$H_o: \phi = 0$$

$$H_a: \phi \neq 0$$

on the first two sample autocorrelations r_1 and r_2 (i.e., K=2). Call this the case (J=2, K=2). Applying formulas (3.3.9) and (3.3.10) from Theorem 5, we obtain

$$v_{11} = \frac{1+2\phi+\phi^2}{1-\phi^2} = \frac{(1+\phi)^2}{1-\phi^2} ,$$

$$v_{22} = \frac{1-2\phi+\phi^2}{1-\phi^2} = \frac{(1-\phi^2)^2}{1-\phi^2} = 1-\phi^2,$$

and

$$v_{12} = 0.$$

From equation (3.3.35), we obtain $\rho_1 = 0$ and $\rho_2 = \phi$. Thus, we have that, as n tends to infinity, the vector $(\sqrt{n}\, r_1, \sqrt{n}\, r_2 - \sqrt{n}\, \phi)'$ is distributed asymptotically normal with mean vector $(0,0)'$ and dispersion matrix

$$V = \begin{bmatrix} \frac{(1+\phi)^2}{1-\phi^2} & 0 \\ 0 & 1-\phi^2 \end{bmatrix} .$$

We write:

$$(\sqrt{n}\, r_1, \sqrt{n}\, r_2)' \overset{\cdot}{\sim} N \left\{ (0, \sqrt{n}\, \phi)' , \begin{bmatrix} \frac{(1+\phi)^2}{1-\phi^2} & 0 \\ 0 & 1-\phi^2 \end{bmatrix} \right\} , \qquad (3.3.14)$$

for large n.

Corollary 1: Under the alternative error model given in equation (3.3.13), the Box-Pierce statistic $T^{(BP)} = nr_1^2 + nr_2^2$, is asymptotically distributed as a weighted sum of independent single degree of freedom chi-square random variables:

$$T^{(BP)} = \frac{(1+\phi)^2}{1-\phi^2} \cdot \chi_1^2(0) + (1-\phi)^2 \cdot \chi_1^2(\lambda) ,$$

where $\lambda = \dfrac{n^2 \phi^2}{1-\phi^2}$.

Proof: From Theorem 2, we write $T^{(BP)} = \sum_{i=1}^{\ell} \lambda_i \chi_{m_i}^2(\delta_i^2)$, where the $\{\lambda_i\}$, $\{m_i\}$, and $\{\delta_i^2\}$ are defined previously. Since V is a diagonal matrix, it is obvious that the eigenvalues of V are:

$$\lambda_1 = \frac{(1+\phi)^2}{1-\phi^2} \quad \text{and} \quad \lambda_2 = 1-\phi^2 ,$$

with corresponding multiplicities $m_1=1$ and $m_2=1$. In order to determine δ_1^2 and δ_2^2 we utilize (iv) of Theorem 2. Now

$$E_1 = \left(\frac{1}{\lambda_1 - \lambda_2} \right) (T'T - \lambda_2 \cdot I)$$

and

$$E_2 = \left(\frac{1}{\lambda_2 - \lambda_1} \right) (T'T - \lambda_1 \cdot I) ,$$

where

$$T = \begin{bmatrix} \frac{1+\phi}{\sqrt{1-\phi^2}} & 0 \\ 0 & \sqrt{1-\phi^2} \end{bmatrix} \qquad \text{(Recall that } V=TT'\text{)}.$$

Hence, we have

$$\delta_1^2 = n \cdot \underset{\sim}{\rho}' V^{-1} T E_1 T^{-1} \underset{\sim}{\rho}$$

$$= n \cdot \left(0, \frac{\sqrt{n}\ \phi}{1-\phi^2}\right)' T E_1 T^{-1} \underset{\sim}{\rho}$$

$$= n \cdot \left(0, \frac{\sqrt{n}\ \phi}{(1-\phi^2)^{3/2}}\right)' E_1 T^{-1} \underset{\sim}{\rho}$$

$$= n \cdot \left(0, \frac{\sqrt{n}\ \phi}{(1-\phi^2)^{3/2}} \cdot \left[\frac{1-\phi^2-\lambda_2}{\lambda_1-\lambda_2}\right]\right)' T^{-1} \underset{\sim}{\rho}$$

$$= n \cdot \left(0, \frac{\sqrt{n}\ \phi}{(1-\phi^2)^2} \left[\frac{1-\phi^2-\lambda_2}{\lambda_1-\lambda_2}\right]\right)' \underset{\sim}{\rho}$$

$$= \frac{n^2\phi^2[1-\phi^2-\lambda_2]}{(1-\phi^2)^2(\lambda_1-\lambda_2)} = 0 \ ,$$

since $\lambda_2 = 1-\phi^2$. Similarly,

$$\delta_2^2 = n \cdot \underset{\sim}{\rho} V^{-1} T E_2 T^{-1} \underset{\sim}{\rho}$$

$$= \frac{n^2\phi^2[1-\phi^2-\lambda_1]}{(1-\phi^2)^2(\lambda_2-\lambda_1)}$$

$$= \frac{n^2\phi^2}{(1-\phi^2)^2} \ . \qquad \square$$

From Corollary 1, then, we write the exact asymptotic power of the Box-Pierce test as:

$$\lim_{n\to\infty} P\{T^{(BP)} > C_\alpha^{(BP)} \mid \phi\} = P\left\{\frac{(1+\phi)^2}{1-\phi^2}\ \chi_1^2(0) + (1-\phi^2)\chi_1^2(\lambda) > C_\alpha^{(BP)}\right\} \ , \qquad (3.3.15)$$

where $\lambda = n^2\phi^2/(1-\phi^2)^2$. Now the probability on the right-hand side of equation (3.3.15) is not an easy one to compute, since the constants mulitplying the single degree of freedom χ^2 random variables are unequal.

In fact, the problem of obtaining the exact probability remains unsolved. In Section 3.3.3.2, we present three methods of approximating the right hand side of (3.3.15).

Corollary 2: Under the alternative error model given in equation (3.3.13), the exact asymptotic power of the max-χ^2 test, $T^{(m)} = \max(nr_1^2, nr_2^2)$, is given by:

$$\lim_{n\to\infty} P\{T^{(m)} > C_\alpha^{(m)}\} = 1 - \left[2\cdot\Phi\left\{\frac{\sqrt{C_\alpha^{(m)}(1-\emptyset^2)}}{1+\emptyset}\right\} - 1\right]\cdot\left[\Phi\left\{\frac{\sqrt{C_\alpha^{(m)}} - \sqrt{n}\emptyset}{\sqrt{1-\emptyset^2}}\right\} - \Phi\left\{\frac{-\sqrt{C_\alpha^{(m)}} - \sqrt{n}\emptyset}{\sqrt{1-\emptyset^2}}\right\}\right], \tag{3.3.16}$$

where $\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}z^2}\, dz$

Proof: Let $C_\alpha^{(m)} = C$ and start by writing the power of the max-χ^2 test as follows:

$$\begin{aligned} P\{T^{(m)} > C \mid \emptyset\} &= P\{\max(nr_1^2, nr_2^2) > C\} \\ &= 1 - P\{\max(nr_1^2, nr_2^2) < C\} \\ &= 1 - P\{nr_1^2 < C,\ nr_2^2 < C\}\ . \end{aligned}$$

From the asymptotic distribution given in (3.3.14), we have that for large n, $\sqrt{n}r_1$ and $\sqrt{n}r_2$ are distributed as independent normal random variates; hence,

$$\begin{aligned} \lim_{n\to\infty} P\{T^{(m)} > C \mid \emptyset\} &= 1 - P\{nr_1^2 < C\}\cdot P\{nr_2^2 < C\} \\ &= 1 - P\{-\sqrt{C} < \sqrt{n}r_1 < \sqrt{C}\}\cdot P\{-\sqrt{C} < \sqrt{n}r_2 < \sqrt{C}\} \end{aligned}$$

$$= 1-P\left\{\frac{-\sqrt{C}}{\frac{(1+\emptyset)^2}{1-\emptyset^2}} < Z < \frac{\sqrt{C}}{\frac{(1+\emptyset)^2}{1-\emptyset^2}}\right\} \cdot P\left\{\frac{-\sqrt{C}-\sqrt{n}\emptyset}{\sqrt{1-\emptyset^2}} < Z < \frac{\sqrt{C}-\sqrt{n}\emptyset}{\sqrt{1-\emptyset^2}}\right\}$$

$$= 1-\left[\Phi\left\{\frac{\sqrt{C(1-\emptyset^2)}}{1+\emptyset}\right\} - \Phi\left\{\frac{-\sqrt{C(1-\emptyset^2)}}{1+\emptyset}\right\}\right]\cdot\left[\Phi\left\{\frac{\sqrt{C}-\sqrt{n}\emptyset}{\sqrt{1-\emptyset^2}}\right\} - \Phi\left\{\frac{-\sqrt{C}-\sqrt{n}\emptyset}{\sqrt{1-\emptyset^2}}\right\}\right]$$

$$= 1-\left[2\cdot\Phi\left\{\frac{\sqrt{C(1-\emptyset^2)}}{1+\emptyset}\right\}-1\right]\cdot\left[\Phi\left\{\frac{\sqrt{C}-\sqrt{n}\emptyset}{\sqrt{1-\emptyset^2}}\right\} - \Phi\left\{\frac{-\sqrt{C}-\sqrt{n}\emptyset}{\sqrt{1-\emptyset^2}}\right\}\right] . \quad \square$$

Of course, tables are available for computing tail areas of the standard normal distribution. Thus, for the special case (J=2, K=2), the exact asymptotic power can be calculated for the max-χ^2 test. However, the notion of exact asymptotic power is the exception rather than the rule. The researcher who attempts to calculate powers utilizing an analytic approach similar to the work above, will more often than not be forced into approximating the true asymptotic power, as with the Box-Pierce statistic. This point is further illustrated in the following section.

3.3.3.2 Model 2: The Case J=2, K=4

Again, let us consider model (3.3.13) as an appropriate representation of the residuals when serial correlation is present. Now we base the test on the first four sample autocorrelations (r_1, r_2, r_3, r_4). Call this the case (J=2, K=4). From equation (3.3.5), we observe that $\rho_1=0$, $\rho_2=\emptyset$, $\rho_3=0$ and $\rho_4=\emptyset^2$. Applying (3.3.9) gives

$$v_{11} = \frac{(1+\emptyset)^2}{1-\emptyset^2} \quad ,$$

$$v_{22} = 1 - \emptyset^2 \quad ,$$

$$v_{33} = \frac{1+\emptyset^2+4\emptyset^3-2\emptyset^5}{1-\emptyset^2} \quad ,$$

$$v_{44} = \frac{1+\phi^2-5\phi^4+3\phi^6}{1-\phi^2} = 1+2\phi^2-3\phi^4 \ .$$

The covariances, obtained from (3.3.10), are computed as follows:

$$v_{12} = 0 \ ,$$

$$v_{13} = \frac{\phi^2(3-\phi^2)+2\phi}{1-\phi^2} = \frac{2\phi+3\phi^2-\phi^4}{1-\phi^2} \ ,$$

$$v_{14} = 0,$$

$$v_{23} = 0,$$

$$v_{24} = \frac{2\phi-\phi^3(4-2\phi^2)}{1-\phi^2} = \frac{2\phi(1-2\phi^2+\phi^4)}{1-\phi^2} = 2\phi(1-\phi^2) \ .$$

Then, as n tends to infinity, the vector $\sqrt{n}(\underset{\sim}{r}-\underset{\sim}{\rho})$ is distributed asymptotically normal, with mean vector $\underset{\sim}{0}$, and dispersion matrix V, where

$$V = \begin{bmatrix} \frac{(1+\phi)^2}{1-\phi^2} & 0 & \frac{2\phi+3\phi^2-\phi^4}{1-\phi^2} & 0 \\ 0 & 1-\phi^2 & 0 & 2\phi(1-\phi^2) \\ \frac{2\phi+3\phi^2-\phi^4}{1-\phi^2} & 0 & \frac{1+\phi^2+4\phi^3-2\phi^5}{1-\phi^2} & 0 \\ 0 & 2\phi(1-\phi^2) & 0 & 1+2\phi^2-3\phi^4 \end{bmatrix} \qquad (3.3.17)$$

The method of obtaining exact asymptotic powers in the case (J=2, K=2) depended heavily on the independence of the first two sample autocorrelations. The addition of the third and fourth order sample autocorrelations into the testing procedure contribute nonzero off-diagonal elements to the covariance structure, V. These covariances make the distributional problem so complex, that the researcher is forced to abandon the search for exact asymptotic powers and seek, instead, approximate asymptotic powers. In the subsections that follow, we examine several approximate procedures for the Box-Pierce test, and one approximation for the max-χ^2 test.

3.3.3.2.1 Approximate asymptotic power of the Box-Pierce test

Three approximate methods for computing the asymptotic power of the Box-Pierce test are given in theorem form, and are labeled as follows:

(i) Imhoff numerical integration

(ii) Extension to Pearson's three-moment approximation

(iii) Rao asymptotic normality.

Lemma 6 (Imhoff): Let $\underset{\sim}{x}$ have an n-variate normal distribution with mean $\underset{\sim}{\mu}$ and covariance V, where V is an nxn symmetric positive definite matrix. Consider the quadratic form,

$$Q = \underset{\sim}{x}'\Lambda\,\underset{\sim}{x} = \sum_{i=1}^{\ell} \lambda_i \chi^2_{m_i}(\delta_i^2),$$

where $\{\lambda_i\}$, $\{m_i\}$, and $\{\delta_i^2\}$ are defined in Lemma 2. Then

$$P\{Q>C\} = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\sin\theta(u)}{u\,\rho(u)}\,du \quad ,$$

where

$$\theta(u) = \frac{1}{2}\sum_{i=1}^{\ell}[m_i \tan^{-1}(\lambda_i \cdot u)+\delta_i^2\lambda_i u(1+\lambda_i^2 u^2)^{-1}] - \frac{1}{2}Cu \ ,$$

and

$$\rho(u) = \prod_{i=1}^{\ell}(1+\lambda_i^2u^2)^{\frac{1}{4}m_i}\cdot\exp\{\frac{1}{2}\sum_{j=1}^{\ell}(\delta_j\lambda_j u)^2/(1+\lambda_i^2u^2)\} \ .$$

Proof: Imhoff's result is obtained by inversion of the characteristic function of the variable Q. For details, see Imhoff [23], pp.421. ☐

Theorem 7 (Imhoff numerical integration): Under the alternative model given by (3.3.13), the asymptotic power of the Box-Pierce test, based on K=4 sample autocorrelations, is given by

$$\lim_{n\to\infty} P\{ T^{(BP)} > C_\alpha^{(BP)} \} = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\sin\,\theta(u)}{u\rho(u)}\,du\ , \qquad (3.3.18)$$

where

$$\theta(u) = \frac{1}{2}\sum_{i=1}^{\ell}[m_i \tan^{-1}(\lambda_i u) + \delta_i^2\lambda_i u(1+\lambda_i^2u^2)^{-1}] - \frac{1}{2}C_\alpha^{(BP)}u\ ,$$

$$\rho(u) = \prod_{i=1}^{\ell}(1+\lambda_i^2u^2)^{\frac{1}{4}m_i}\exp\{\frac{1}{2}\sum_{j=1}^{\ell}(\delta_j\lambda_j u)^2/(1+\lambda_i^2u^2)\ ,$$

and where $\{\lambda_i\}$ are the eigenvalues of the matrix V given in (3.3.17), with respective multiplicities $\{m_i\}$, and $\{\delta_i^2\}$ are obtained from (iv) of Theorem 2.

Proof: The result follows directly from Theorem 2 and Lemma 6. □

The power expression in (3.3.18) results from Imhoff's ability to find the distribution function of a linear combination of non-central χ^2 random variables explicitly. However, the integral in (3.3.18) cannot be computed analytically. Imhoff outlines a procedure in which the researcher can approximate the power by a numerical integration method. In Chapter 4, approximate asymptotic powers of the Box-Pierce test for this case (J=2, K=4) and others are calculated by means of Imhoff's procedure using numerical integration.

Lemma 7 (Pearson): Given the conditions of Lemma 3, tail areas of the distribution of the quadratic form Q may be approximated as follows:

$$P\{ Q > C\} \approx P\{ \chi_h^2(0) > C^*\},$$

where $h = a_2^3/a_3^2$, $C^* = (C-a_1)\cdot(h/a_2)^{\frac{1}{2}}+h$, $a_j = \sum_{i=1}^{\ell}\lambda_i^j(m_i+j\delta_i^2)$ $(j=1,2,3)$ and $\chi_h^2(0)$ a central χ^2 random variable with h degrees of freedom.

Proof: The result is derived by extending Pearson's [25] "three-moment central χ^2 approximation to the distribution of a non-central χ^2" to the general case of a quadratic form in normal random variables. Details are given in Imhoff [23], pp. 425. □

Theorem 7 (Imhoff's extension to Pearson 3-moment approximation): An approximation to the asymptotic power of the Box-Pierce test is given by:

$$\lim_{n\to\infty} P\{T^{(BP)} > C_\alpha^{(BP)}\} \approx P\{\chi_h^2(0) > C^*\}, \qquad (3.3.19)$$

where $h = a_2^3/a_3^2$, $C^* = (C_\alpha^{(BP)} - a_1)\cdot(h/a_2)^{\frac{1}{2}} + h$, $a_j = \sum_{i=1}^{\ell} \lambda_i^j (m_i + j\delta_i^2)$, $(j=1,2,3)$, $\{\lambda_i\}$, $\{m_i\}$, and $\{\delta_i^2\}$ defined as in (iv) of Theorem 2, and where "$\approx$" represents Pearson's 3-moment central χ^2 approximation to the distribution of a non-central χ^2.

Proof: The results of Theorem 2 enable us to write $T^{(BP)}$ as a quadratic form, Q. Apply Imhoff's extension to Pearson's 3-moment approximation (Lemma 7) and the result follows. ☐

Since h in Theorem 7 need not necessarily be an integer, the user of this extended three-moment approximate method must have access to a computer routine which computes tail areas of central χ^2 random variables with non-integer degrees of freedom. For a numerical application of this result, see Section (4.1).

Lemma 8 (Rao): Let $\sqrt{n}(\underset{\sim}{x} - \underset{\sim}{\mu}) \xrightarrow{D} N_K\{0, \Sigma\}$ and let $g: R^K \longrightarrow R$ be a continuous function. Then, under certain regularity conditions,

$$\sqrt{n}\{g(\underset{\sim}{x}) - g(\underset{\sim}{\mu})\} \xrightarrow{D} N\{0, \underset{\sim}{\zeta}'\Sigma\underset{\sim}{\zeta}\},$$

where

$$\underset{\sim}{\zeta} = \left(\frac{\partial g(\underset{\sim}{t})}{\partial t_1}\bigg|_{\mu_1}, \frac{\partial g(\underset{\sim}{t})}{\partial t_2}\bigg|_{\mu_2}, \ldots, \frac{\partial g(\underset{\sim}{t})}{\partial t_K}\bigg|_{\mu_K}\right)'.$$

Proof: See C.R. Rao [27], pp. 387. ☐

Theorem 8: (Rao's asymptotic normal tranformation): An alternate form for the asymptotic power of the Box-Pierce test is given by:

$$\lim_{n\to\infty} P\{T^{(BP)}>C_\alpha^{(BP)}\} = 1-\Phi\left\{\frac{\sqrt{C_\alpha^{(BP)}}- n(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}}\right\} +\Phi\left\{\frac{-\sqrt{C_\alpha^{(PB)}}-\sqrt{n}(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}}\right\}, \quad (3.3.20)$$

where

$$\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-\frac{1}{2}}e^{-z^2/2}dz ,$$

where $\underset{\sim}{\rho}$ and V are the asymptotic mean vector and asymptotic dispersion matrix, respectively, of the vector of sample autocorrelations, r.

Proof: From Lemma 1, we have $\sqrt{n}(\underset{\sim}{r}-\underset{\sim}{\rho}) \xrightarrow{D} N_K\{\underset{\sim}{0},V\}$; hence, we may apply the result of Lemma 8. Consider the function

$$g(\underset{\sim}{t}) = (\underset{\sim}{t}'\underset{\sim}{t})^{\frac{1}{2}} =\sqrt{t_1^2 + t_2^2 + \ . \ . \ . + t_K^2} \ .$$

Then,

$$\frac{\partial g(\underset{\sim}{t})}{\partial t_i} = \frac{1}{2}(t_1^2 + t_2^2 + \ . \ . \ . + t_K^2)^{-\frac{1}{2}}(2t_i)$$

$$= t_i(\underset{\sim}{t}'\underset{\sim}{t})^{-\frac{1}{2}} \ , \quad (i = 1,\ 2,\ .\ .\ .,\ K),$$

and

$$\left.\frac{\partial g(\underset{\sim}{t})}{\partial t_i}\right|_{\rho_i} = \rho_i(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}} \ , \quad (i = 1,\ 2,\ .\ .\ .,\ K),$$

imply that $\underset{\sim}{\zeta} = (\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}\underset{\sim}{\rho}$. Thus, from Lemma 8:

$$\sqrt{n}\{(\underset{\sim}{r}'\underset{\sim}{r})^{\frac{1}{2}} - (\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}\}\xrightarrow{D} N\{0,\ (\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-1}\underset{\sim}{\rho}'V\underset{\sim}{\rho}\} \ .$$

For large n, we see that $\sqrt{n}(\underset{\sim}{r}'\underset{\sim}{r})^{\frac{1}{2}}$ has an approximate normal distribution with mean $\sqrt{n}(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}$ and variance $(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-1}\underset{\sim}{\rho}'V\underset{\sim}{\rho}$. We compute the asymptotic power by writing:

$$P\{T^{(BP)}>C\} = P\{n\underset{\sim}{r}'\underset{\sim}{r}>C\} = 1-P\{n\underset{\sim}{r}'\underset{\sim}{r}<C\}$$

$$= 1-P\{-\sqrt{C}<\sqrt{n}(\underset{\sim}{r}'\underset{\sim}{r})^{\frac{1}{2}}<\sqrt{C}\} \ .$$

Taking limits on both sides, we have

$$\lim_{n\to\infty} P\{T^{(BP)}>C\} = 1-P\left\{\frac{-\sqrt{C}-\sqrt{n}(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}} < Z < \frac{\sqrt{C}-\sqrt{n}(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}}\right\}$$

$$= 1-\Phi\left\{\frac{\sqrt{C}-\sqrt{n}(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}}\right\} + \Phi\left\{\frac{-\sqrt{C}-\sqrt{n}(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}}\right\} \quad . \quad \square$$

We apply this result (Theorem 8) to the case (J=2, K=4).

Corollary 3: For the case (J=2, K=4), the asymptotic power of the Box-Pierce test is given by the expression:

$$\lim_{n\to\infty} P\{T^{(BP)}>C_\alpha^{(BP)}\} = 1-\Phi\left\{\frac{\sqrt{C_\alpha^{(BP)}}-\phi\sqrt{n(1+\phi^2)}}{\frac{\sqrt{1+4\phi^2-2\phi^4-3\phi^6}}{1+\phi^2}}\right\} + \Phi\left\{\frac{-\sqrt{C_\alpha^{(BP)}}-\phi\sqrt{n(1+\phi^2)}}{\frac{\sqrt{1+4\phi^2-2\phi^4-3\phi^6}}{1+\phi^2}}\right\} . \quad (3.3.21)$$

Proof: Recall that, for (J=2, K=4), $\underset{\sim}{\rho} = (0, \phi, 0, \phi^2)'$, and the matrix V is given by (3.3.17). Then $\underset{\sim}{\rho}'\underset{\sim}{\rho} = \phi^2 + \phi^4 = \phi^2(1+\phi^2)$ and

$$\rho'V\rho = [0,\ \phi(1-\phi^2)+\phi^2[2\phi(1-\phi^2)],\ 0,\ \phi[2\phi(1-\phi^2)]+\phi^2(1-2\phi^2-3\phi^4)] . \begin{bmatrix} 0 \\ \phi \\ 0 \\ \phi^2 \end{bmatrix}$$

$$= \phi^2(1-\phi^2)+2\phi^4(1-\phi^2)+2\phi^4(1-\phi^2)+\phi^4(1+2\phi^2-3\phi^4)$$

$$= \phi^2+4\phi^4-2\phi^6-3\phi^8$$

$$= \phi^2(1+4\phi^2-2\phi^4-3\phi^6) .$$

Thus,

$$(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}} = [\phi\sqrt{1+\phi^2}]^{-1} \cdot \phi \cdot \sqrt{1+4\phi^2-2\phi^4-3\phi^6}$$

$$= \sqrt{\frac{1+4\phi^2-2\phi^4-3\phi^6}{1+\phi}}$$

Now utilize equation (3.3.20) in Theorem 8 to obtain the approximate asymptotic power expression:

$$1-\Phi\left\{\frac{\sqrt{C_\alpha^{(BP)}} - \phi\sqrt{n(1+\phi^2)}}{\sqrt{\frac{1+4\phi^2-2\phi^4-3\phi^6}{1+\phi^2}}}\right\} + \Phi\left\{\frac{-\sqrt{C_\alpha^{(BP)}} - \phi\sqrt{n(1+\phi^2)}}{\sqrt{\frac{1+4\phi^2-2\phi^4-3\phi^6}{1+\phi^2}}}\right\} \quad . \quad \square$$

A comparison of the three approximations for computing the power of the Box-Pierce statistic is found in Chapter 4.

3.3.3.2.2 Approximate asymptotic power of the max-χ^2 test.

We now restrict our attention to the max-χ^2 test. For the case (J=2, K=4), the asymptotic power of $T^{(m)}$ may be written:

$$P\{T^{(m)}>C\} = P\{n \max_{1\le j\le 4} r_j^2 > C\}$$

$$= 1-P\{nr_1^2<C,\ nr_2^2<C,\ nr_3^2<C,\ nr_4^2<C\}$$

$$= 1-P\{-\sqrt{C}<\sqrt{n}r_1<\sqrt{C},\ .\ .\ .,\ -\sqrt{C}<\sqrt{n}r_4<\sqrt{C}\}$$

Let $x_i=\sqrt{n}r_i$ and $\mu_i= n\rho_i$, $1\le i\le 4$. Then the problem is to evaluate

$$\int_{-\sqrt{C}}^{\sqrt{C}}\int_{-\sqrt{C}}^{\sqrt{C}}\int_{-\sqrt{C}}^{\sqrt{C}}\int_{-\sqrt{C}}^{\sqrt{C}} f(\underset{\sim}{x}:\ \underset{\sim}{\mu},\ V)\ dx_1dx_2dx_3dx_4, \qquad (3.3.22)$$

where $f(\underset{\sim}{x}:\ \underset{\sim}{\mu},V)$ is the joint density function of the first four sample autocorrelations. For large n, $\sqrt{n}\underset{\sim}{r}$ is multivariate normal; thus

$$f(\underset{\sim}{x}:\ \underset{\sim}{\mu},V) = \frac{|V|^{-\frac{1}{2}}}{(2\pi)^2}\exp\{-\ \frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})'V^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})\}\ .$$

Due to the presence of offdiagonal elements in the matrix V, the exact computation of the multiple integral in (3.3.22) becomes extremely cumbersome. The researcher must again resort to approximate methods. We propose a Taylor-series expansion method of approximating (3.3.22).

Definition 1: Let $\underset{\sim}{y} = (y_1,\ y_2,\ .\ .\ .,\ y_K)'$ be distributed as multivariate normal, with $E(y_i) = 0$, $E(y_i^{\ 2}) = 1$, $1\le i\le K$, and $E(y_iy_j)=C_{ij}$, $1\le i\le K$, $1\le j\le K$, $i\ne j$ (i.e., $\underset{\sim}{y} \sim N_K\{\underset{\sim}{0},\ R\}$, where $R=(C_{ij})$ is a correlation matrix). A Taylor-series expansion of the density of $\underset{\sim}{y}$, $f(\underset{\sim}{y}:\ \underset{\sim}{0},R)$, about the elements C_{ij} in R is defined as follows:

$$f(\underset{\sim}{y}: \underset{\sim}{0},R) = \sum_{t=0}^{\infty}\left[\sum_{\substack{i=1\\ i<j}}^{K}\sum_{j=1}^{K}\frac{(C_{ij})^t}{t!}\cdot\left(\frac{\partial}{\partial C_{ij}}\right)^t f(\underset{\sim}{y}: \underset{\sim}{0},R)\Big|_{\{C_{ij}\}=0}\right] \qquad (3.3.23)$$

where $\{C_{ij}\}=0$ implies $C_{ij}=0\ \forall i,j\ (i\neq j)$. $\square$

For notational purposes, let

$$\left(\frac{\partial}{\partial C_{ij}}\right)^t f(\underset{\sim}{y}: \underset{\sim}{0},R)\Big|_{\{C_{ij}\}=0} = \left(\frac{\partial}{\partial C_{ij}}\right)^t f_o.$$

Expanding (3.3.23) we obtain:

$$f(\underset{\sim}{y}: \underset{\sim}{0},R) = \{f(\underset{\sim}{y}: \underset{\sim}{0},I)\} + C_{12}\frac{\partial}{\partial C_{12}}f_o + C_{13}\frac{\partial}{\partial C_{13}}f_o + \dots \qquad (3.3.24)$$

$$+ C_{K-1,K}\frac{\partial}{C_{K-1,K}}f_o$$

$$+\frac{1}{2}\{C_{12}^2\left(\frac{\partial}{C_{12}}\right) f_o + \dots + C_{K-1,K}^2\left(\frac{\partial}{\partial C_{K-1,K}}\right)^2 f_o + C_{12}C_{13}\left(\frac{\partial}{C_{12}C_{13}}\right)f_o$$

$$+ \dots C_{K-2,K}C_{K-1,K}\left(\frac{\partial}{\partial C_{K-2,K}C_{K-1,K}}\right)f_o\}$$

$$+\frac{1}{3!}\{C_{12}^3\left(\frac{\partial}{\partial C_{12}}\right)^3 f_o + \dots + C_{K-1,K}^3\left(\frac{\partial}{\partial C_{K-1,K}}\right)^3 f_o + C_{12}^2C_{13}\left(\frac{\partial}{\partial C_{12}^2\partial C_{13}}\right)f_o$$

$$+ \dots C_{K-2,K}C_{K-1,K}^2\left(\frac{\partial}{\partial C_{K-2,K}C_{K-1,K}^2}\right)f_o\}$$

$$+\frac{1}{4!}\{\ \dots\ \} + \dots \qquad .$$

We will attempt to simplify equation (3.3.24) in order to make it more workable for the practical researcher. To do this, we introduce the concept of Hermite polynomials.

Definition 2: Let $\phi(x)$ represent the density function of a univariate $N(0,1)$ random variable (i.e., $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$). The n^{th}

Hermite polynomial, $H_n(x)$, is defined from the following identity:

$$(-1)^n H_n(x)\,\phi(x) = \left(\frac{\partial}{\partial x}\right)^n \phi(x). \qquad \square \qquad (3.3.25)$$

The properties of Hermite polynomials are given in the following lemma.

Lemma 9: Let $H_n(x)$ be the n^{th} Hermite polynomial. Then $H_o(x)=1$, $H_1(x)=x$, $H_2(x)=(x^2-1)$, $H_3(x)=(x^3-3x)$ and, in general,

$$H_{n+1}(x) = xH_n(x) - \left(\frac{\partial}{\partial x}\right)H_n(x) \quad . \qquad (3.3.26)$$

Proof: By convention, take $H_o(x)=1$. Now,

$$\left(\frac{\partial}{\partial x}\right)\phi(x) = -x\phi(x),$$

$$\left(\frac{\partial}{\partial x}\right)^2 \phi(x) = -\phi(x) + x^2\phi(x) = (x^2-1)\phi(x) \quad ,$$

and

$$\left(\frac{\partial}{\partial x}\right)^3 \phi(x) = (2x)\phi(x) - x(x^2-1)\phi(x) = -(x^3-3x)\phi(x).$$

By equation 3.3.25, then, $H_1(x)=x$, $H_2(x)=(x^2-1)$ and $H_3(x)=(x^3-3x)$. In general, the $(n+1)^{st}$ Hermite polynomial is found as follows:

$$\begin{aligned}(-1)^{n+1}H_{n+1}(x)\phi(x) &= \left(\frac{\partial}{\partial x}\right)^{n+1}\phi(x)\\ &= \left(\frac{\partial}{\partial x}\right)\left\{\left(\frac{\partial}{\partial x}\right)^n\phi(x)\right\}\\ &= \left(\frac{\partial}{\partial x}\right)(-1)^n H_n(x)\phi(x)\\ &= (-1)^n\left\{\left[\left(\frac{\partial}{\partial x}\right)H_n(x)\right]\cdot\phi(x) - xH_n(x)\phi(x)\right\} \quad .\end{aligned}$$

Thus, we obtain the $(n+1)^{st}$ Hermite polynomial from the recursive relation:

$$H_{n+1}(x) = xH_n(x) - \left(\frac{\partial}{\partial x}\right)H_n(x) \quad . \qquad \square$$

In the lemmas that follow, the application of Hermite polynomials in the problem at hand becomes evident.

Lemma 10: Given the conditions of Definition 1, we have:

$$\text{(i)}\quad \left(\frac{\partial}{\partial C_{ij}}\right)^t f(\underset{\sim}{y}: \underset{\sim}{0},R) = \left(\frac{\partial}{\partial y_i \partial y_j}\right)^t f(\underset{\sim}{y}: \underset{\sim}{0},R), \quad 1\le i\le K,\ 1\le j\le K,\ i\ne j$$

$$\text{(ii)}\quad \left(\frac{\partial}{\partial y_i^m \partial y_j^n}\right) f(\underset{\sim}{y}: \underset{\sim}{0},R)\Bigg|_{\{C_{ij}\}=0} = H_m(y_i)H_n(y_j)f(\underset{\sim}{y}: \underset{\sim}{0},I).$$

Proof: The proof of (i) is straightforward. Perform the required partial differentiation of the density $f(\underset{\sim}{y}: \underset{\sim}{0},R)$ on each side of the equation, and the result follows (this requires many tedious computations).

We derive the identity in (ii) for the case K=2. For the case K=2, we have

$$R = \begin{bmatrix} 1 & C_{12} \\ C_{12} & 1 \end{bmatrix}, \qquad R^{-1} = \left(\frac{1}{1-C_{12}^2}\right) \begin{bmatrix} 1 & -C_{12} \\ -C_{12} & 1 \end{bmatrix}$$

and

$$f(\underset{\sim}{y}: \underset{\sim}{0},R) = \frac{|R|^{-\frac{1}{2}}}{(2\pi)} \exp\left\{ -\frac{1}{2}(y_1^2-2C_{12}y_1y_2+y_2^2)\cdot \frac{1}{1-C_{12}^2}\right\}$$

Applying (i) and taking the partial derivatives, we obtain

$$\begin{aligned}\left(\frac{\partial}{\partial C_{12}}\right) f(\underset{\sim}{y}: \underset{\sim}{0},R) &= \left(\frac{\partial}{\partial y_1 \partial y_2}\right) f(\underset{\sim}{y}: \underset{\sim}{0},R) \\ &= \left(\frac{\partial}{\partial y_2}\right)\left\{ f(\underset{\sim}{y}: \underset{\sim}{0},R)\cdot\left(\frac{C_{12}y_2-y_1}{1-C_{12}^2}\right)\right\} \\ &= f(\underset{\sim}{y}: \underset{\sim}{0},R)\left(\frac{C_{12}}{1-C_{12}^2}\right) + f(\underset{\sim}{y}: \underset{\sim}{0},R)\left(\frac{C_{12}y_2-y_1}{1-C_{12}^2}\right)\left(\frac{C_{12}y_1-y_2}{1-C_{12}^2}\right)\end{aligned}$$

Taking $C_{12}=0$ in the above equation gives:

$$\left(\frac{\partial}{\partial C_{12}}\right) f(\underset{\sim}{y}: \underset{\sim}{0},R)\Big|_{\{C_{ij}\}=0} = y_1 y_2 f(\underset{\sim}{y}: \underset{\sim}{0},I) = H_1(y_1)H_1(y_2)f(\underset{\sim}{y}: \underset{\sim}{0},I).$$

The proof of the more general result (ii) follows a similar argument. □

Lemma 11: Given the conditions of Definition 1, the Taylor-series expansion of $f(\underset{\sim}{y}: \underset{\sim}{0},R)$ can be expressed as follows:

$$f(\underset{\sim}{y}: \underset{\sim}{0},R) = \{[1] + [C_{12}H_1(y_1)H_1(y_2) + C_{13}H_1(y_1)H_1(y_3) + \cdots +$$

$$C_{K-1,K}H_1(y_{K-1})H_1(y_K)]$$

$$+ \frac{1}{2}[C_{12}^2 H_2(y_1)H_2(y_2) + \ldots + C_{K-1,K}^2 H_2(y_{K-1})H_2(y_K) + C_{12}C_{13}H_2(y_1)H_1(y_3)H_1(y_2)$$

$$+ \cdots + C_{K-2,K}C_{K-1,K}H_1(y_{K-2})H_1(y_{K-1})H_2(Y_K)]$$

$$+ \frac{1}{3}[C_{12}^3 H_3(y_1)H_3(y_2) + \cdots + C_{K-1,K}^3 H_3(y_{K-1})H_3(y_K)$$

$$+ C_{12}^2 C_{13}H_3(y_1)H_2(y_2)H_1(y_3) + \cdots + C_{K-2,K}C_{K-1}^2 H_1(y_{K-2})H_2(y_{K-1})H_3(y_K) + \cdots$$
$$+ C_{12}C_{13}C_{14}H_3(y_1)H_1(y_2)H_1(y_3)H_1(y_4) + \cdots$$
$$+ C_{K-3,K}C_{K-2,K}C_{K-1,K}H_1(y_{K-2})H_1(y_{K-1})H_3(y_K)H_1(y_{K-3})]$$

$$+ \frac{1}{4!}[\ldots] + \ldots\} f_o$$

where $f_o = f(\underset{\sim}{y}: \underset{\sim}{0},I) = \phi(y_1)\phi(y_2) \ldots \phi(y_K)$.

Proof: Apply the results of Lemma 10 directly to equation (3.3.24). □

Lemma 12:

$$\int_a^b H_n(x)\phi(x)d_x = \begin{cases} \Phi(b)-\Phi(a) & \ldots n=0 \\ H_{n-1}(a)\phi(a)-H_n-1(b)\phi(b) & \ldots n \geq 1 \end{cases}$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

and

$$\Phi(x) = \int_{-\infty}^{x} \phi(z)\,dz \quad .$$

Proof: For n=0,

$$\int_a^b H_o(x)\phi(x)\,dx = \int_a^b \phi(x)\,dx = \int_{-\infty}^{b} \phi(x)\,dx - \int_{-\infty}^{a} \phi(x)\,dx = \Phi(b) - \Phi(a) \; .$$

For $n \geq 1$,

$$\int_a^b H_n(x)\phi(x)\,d_x = \int_a^b d[-H_{n-1}(x)\phi(x)] = -H_{n-1}(x)\phi(x)\Big|_a^b$$

$$= H_{n-1}(a)\phi(a) - H_{n-1}(b)\phi(b) \quad . \qquad \square$$

Lemma 13:

$$\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_K}^{b_K} H_{n_1}(y_1)H_{n_2}(y_2)\cdots H_{n_K}(y_K)\phi(y_1)\phi(y_2)\cdots \phi(y_K)\,d_{y_1}d_{y_2}\cdots d_{y_K}$$

$$= \begin{cases} \prod_{i=1}^{K} [H_{n_i-1}(a_i)\phi(a_i)-H_{n_i-1}(b_i)\phi(b_i)] \, , & n_i \geq 1 \;\forall_i \\ \prod_{i=1}^{K} [\Phi(b_i)-\Phi(a_i)] \, , & n_i = 0 \;\forall_i \\ \prod_{i=1}^{m} [\Phi(b_i)-\Phi(a_i)] \prod_{i=m+1}^{K} [H_{n_i-1}(a_i)\phi(a_i)-H_{n_i-1}(b_i)\phi(b_i)] \, , & \end{cases}$$

where $n_1, n_2, \ldots, n_K$ are reordered so that $n_i=0$, $1 \leq i \leq m$
$n_i \geq 1$, $m+1 \leq i \leq K$.

Proof: Break up the integration as follows:

$$\int_{a_1}^{b_1} H_{n_1}(y_1)\phi(y_1)\left[\int_{a_2}^{b_2} H_{n_2}(y_2)\phi(y_2)\left[\cdots\left[\int_{a_K}^{b_K} H_{n_K}(y_K)\phi(y_K)\,d_{y_K}\right]\cdots\right]\right.$$

$$d_{y_2}\Bigg]\, d_{y_1} \quad .$$

The result now follows directly from Lemma 12. □

We now return to the problem of evaluating the multiple integral in (3.3.22). Let

$$y_i = \frac{x_i - \mu_i}{\sqrt{v_{ii}}}$$

where $V=(v_{ij})$, $1 \leq i \leq 4$, $1 \leq j \leq 4$. Then $\underset{\sim}{y} = (y_1, y_2, y_3, y_4)'$ is multivariate normal with mean $\underset{\sim}{0}$ and dispersion matrix $\underset{\sim}{R}=(C_{ij})$, where $C_{ii}=1$ and

$$C_{ij} = \frac{v_{ij}}{\sqrt{v_{ii} v_{jj}}} \quad .$$

The asymptotic power of the max-χ^2 test for the case (J=2, K=4) may now be written:

$$P\{T^{(m)} > C_\alpha^{(m)}\} = 1 - \int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3}\int_{a_4}^{b_4} f(\underset{\sim}{y}: \underset{\sim}{0}, R)\, dy_1 dy_2 dy_3 dy_4, \quad (3.3.28)$$

where

$$a_i = \frac{-\sqrt{C_\alpha^{(m)}} - \sqrt{n}\rho_i}{\sqrt{v_{ii}}}$$

and

$$b_i = \frac{\sqrt{C_\alpha^{(m)}} - \sqrt{n}\rho_i}{\sqrt{v_{ii}}} \quad , \qquad 1 \leq i \leq 4.$$

The Taylor-series expansion of $f(\underset{\sim}{y}: \underset{\sim}{0}, R)$ is used to obtain an approximation of (3.3.28).

Theorem 9 (Taylor-series approximation): A Taylor-series approximation to the asymptotic power of the max-χ^2 test for the case (J=2,K=4) is obtained by computing

$$1-\Bigg\{[2\Phi(b_1)-1]\cdot[\Phi(b_2)-\Phi(a_2)][2\Phi(b_3)-1][\Phi(b_4)-\Phi(a_4)] \qquad (3.3.29)$$

$$+\{C_{24}[2\Phi(b_1)-1][2\Phi(b_3)-1]\cdot[\phi(a_2)-\phi(b_2][\phi(a_4)-\phi(b_4]\}$$

$$+\frac{1}{2}\{[\Phi(b_2)-\Phi(a_2)][\Phi(b_4)-\Phi(a_4)][C_{13}^2][4b_1\phi(b_1)b_3\phi(b_3)]$$

$$+[2\Phi(b_1)-1][2\Phi(b_3)-1][C_{24}^2][a_2\phi(a_2)-b_2\phi(b_2)][a_4\phi(a_4)-b_4\phi(b_4)]\}$$

$$+\frac{1}{3!}\{[C_{24}^3][2\Phi(b_1)-1][2\Phi(b_3)-1][(a_2^2-1)\phi(a_2)-(b_2^2-1)\phi(b_2)]$$

$$[(a_4^2-1)\phi(a_4)-(b_4^2-1)\phi(b_4)]$$

$$+[C_{13}^2C_{24}][4b_1\phi(b_1)b_3\phi(b_3)][\phi(a_2)-\phi(b_2)][\phi(a_4)-\phi(b_4)]\}$$

$$+\frac{1}{4!}\{[\Phi(b_2)-\Phi(a_2)][\Phi(b_4)-\Phi(a_4)][C_{13}^4][4(b_1^3-3b_1)\phi(b_1)(b_3-3b_3)\phi(b_3)]$$

$$+[C_{24}^4][2\Phi(b_1)-1][2\Phi(b_3)-]][(a_2^3-3a2)\phi(a_2)-(b_2^3-3b_2)\phi(b_2)]\cdot[(a_4^3-3a_4)\phi(a_4)$$

$$-(b_3^3-3b_4)\phi(b_4)]$$

$$+[C_{13}^2C_{24}^2][4b_1\phi(b_1)b_3\phi(b_3)][a_2\phi(a_2)-b_2\phi(b_2)][a_4\phi(a_4)-b_4\phi(b_4)]\}$$

$$+\frac{1}{5!}\{[C_{24}^5][2\Phi(b_1)-1][2\Phi(b_3)-1][(a_2^4-6a_2^2+3)\phi(a_2)-(b_2^4-6b_2^2+3)\phi(b_2)]$$

$$[(a_4^4-6a_4^2+3)\phi(a_4)-(b_4^4-6b_4^2+3)\phi(b_4)]$$

$$+[C_{13}^4C_{24}][4(b_1^3-3b_1)\phi(b_1)(b_3^3-3b_3)\phi(b_3)][\phi(a_2)-\phi(b_2)][\phi(a_4)-\phi(b_4)]$$

$$+[C_{13}^2C_{24}^3][4b_1\phi(b_1)b_3\phi(b_3)][(a_2^2-1)\phi(a_2)-(b_2^2-1)\phi(b_2)][(a_4^2-1)\phi(a_4)$$

$$-(b_4-1)\phi(b_4)]\}$$

$$+\frac{1}{6!}\{[C_{13}^6][\Phi(b_2)-\Phi(a_2)][\Phi(b_4)-\Phi(a_4)][4(b_1^5-10b_1^3+15b_1)\phi(b_1)(b_3^5-10b_3^3+15b_3)$$

$$\phi(b_3)]$$

$$+[C_{13}^4C_{24}^2][4(b_1^3-3b_1)\phi(b_1)(b_3^3-3b_3)\phi(b_3)][a_2\phi(a_2)-b_2\phi(b_2)][a_4\phi(a_4)-b_4\phi(b_4)]$$

$$+[C_{13}^2C_{24}^4][4b_1\phi(b_1)b_3\phi(b_3)][(a_2^3-3a_2)\phi(a_2)-(b_2^3-3b_2)\phi(b_2)]$$

$$[(a_4^3-3a_4)\phi(a_4)-(b_4^3-3b_4)\phi(b_4)]$$

$$+ [C_{24}^{6}][2\Phi(a_1-1)-1][2\Phi(a_3)-1][(a_2^5-10a_2^3+15a_2)\phi(a_2)-(b_2^5-10b_2^3+15b_2)\phi(b_2)]\cdot$$
$$\left.[(a_4^5-10a_4^3+15a_4)\phi(a_4)-(b_4^5-10b_4^3+15b_4)\phi(b_4)]\right\},$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}x^2}, \qquad \Phi(x) = \int_{-\infty}^{x}\phi(z)\,dz,$$

$$C_{13} = \frac{2\varnothing+3\varnothing^2-\varnothing^4}{(1+\varnothing)\sqrt{1+\varnothing^2+4\varnothing^3-2\varnothing^5}}, \quad C_{24} = \frac{2\varnothing\sqrt{1-\varnothing^2}}{\sqrt{1+2\varnothing^2-3\varnothing^4}}, \quad b_1 = \frac{\sqrt{C_\alpha^{(m)}(1-\varnothing^2)}}{1+\varnothing},$$

$$b_2 = \frac{\sqrt{C_\alpha^{(m)}}-\sqrt{n}\,\varnothing}{\sqrt{1-\varnothing^2}}, \quad a_2 = \frac{-\sqrt{C_\alpha^{(m)}}-\sqrt{n}\,\varnothing}{\sqrt{1-\varnothing^2}}, \quad b_3=\frac{\sqrt{C^{(m)}(1-\varnothing^2)}}{\sqrt{1+\varnothing^2+4\varnothing^3-2\varnothing^5}},$$

$$b_4 = \frac{\sqrt{C_\alpha^{(m)}}-\sqrt{n}\,\varnothing}{\sqrt{1+2\varnothing^2-3\varnothing^4}}, \quad \text{and} \quad a_4 = \frac{-\sqrt{C_\alpha^{(m)}}-\sqrt{n}\,\varnothing}{\sqrt{1+2\varnothing^2-3\varnothing^4}}.$$

Proof: From earlier results (3.3.17),

$$v_{11} = \frac{(1+\varnothing)^2}{1-\varnothing^2}, \quad v_{13} = \frac{2\varnothing+3\varnothing^2-\varnothing^4}{1-\varnothing^2},$$

$$v_{22} = 1-\varnothing^2, \quad v_{24} = 2\varnothing(1-\varnothing^2),$$

$$v_{33} = \frac{1+\varnothing^2+4\varnothing^3-2\varnothing^5}{1-\varnothing^2}, \quad v_{44} = 1+2\varnothing^2-3\varnothing^4,$$

$$v_{12} = v_{14} = v_{23} = v_{34} = 0;$$

hence, the elements of R are written,

$$C_{11} = C_{22} = C_{33} = C_{44} = 1,$$

$$C_{12} = C_{14} = C_{23} = C_{34} = 0,$$

$$C_{13} = \frac{2\varnothing+3\varnothing^2-\varnothing^4}{(1+\varnothing)\sqrt{1+\varnothing^2+4\varnothing^3-2\varnothing^5}}, \quad C_{24} = \frac{2\varnothing\sqrt{1-\varnothing^2}}{\sqrt{1+2\varnothing^2-3\varnothing^4}}.$$

Now $\rho_1 = \rho_3 = 0$, $\rho_2 = \phi$, $\rho_4 = \phi^2$ implies that

$$a_1 = \frac{-\sqrt{C_\alpha^{(m)}(1-\phi^2)}}{1+\phi}, \qquad b_1 = \frac{\sqrt{C_\alpha^{(m)}(1-\phi^2)}}{1+\phi},$$

$$a_2 = \frac{-\sqrt{C_\alpha^{(m)}}-\sqrt{n}\,\phi}{\sqrt{1-\phi^2}}, \qquad b_2 = \frac{\sqrt{C_\alpha^{(m)}}-\sqrt{n}\,\phi}{\sqrt{1-\phi^2}},$$

$$a_3 = \frac{\sqrt{C_\alpha^{(m)}(1-\phi^2)}}{\sqrt{1+\phi^2+4\phi^3-2\phi^5}}, \qquad b_3 = \frac{\sqrt{C_\alpha^{(m)}(1-\phi^2)}}{\sqrt{1+\phi^2+4\phi^3-2\phi^5}},$$

$$a_4 = \frac{-\sqrt{C_\alpha^{(m)}}-\sqrt{n}\,\phi^2}{\sqrt{1+2\phi^2-3\phi^4}}, \qquad b_4 = \frac{\sqrt{C_\alpha^{(m)}}-\sqrt{n}\,\phi^2}{\sqrt{1+2\phi^2-3\phi^4}},$$

in the power expression 3.3.28). Put $C_{12} = C_{24} = C_{23} = C_{34} = 0$ in equation (3.3.27), to obtain the Taylor series-expansion:

$$\begin{aligned}
f(\underset{\sim}{y}:\ \underset{\sim}{0},R) = \{&[1]+[C_{13}H_1(y_1)H_1(y_3)+C_{24}H_1(y_2)H_1(y_4)]+\tfrac{1}{2}[C_{13}^2H_2(y_1)H_2(y_3) \qquad (3.3.30)\\
&+ C_{24}^2H_2(y_2)H_2(y_4)+C_{13}C_{24}H_1(y_1)H_1(y_2)H_1(y_3)H_1(y_4)]\\
&+ \tfrac{1}{3!}[C_{13}^3H_3(y_1)H_3(y_3)+ C_{24}^3H_3(y_2)H_3(y_4) + C_{13}^2C_{24}H_2(y_1)H_1(y_2)H_2(y_3)H_1(y_4)\\
&+ C_{13}C_{24}^2H_1(y_1)H_2(y_2)H_1(y_3)H_2(y_4)]\\
&+ \tfrac{1}{4!}[C_{13}^4H_4(y_1)H_4(y_3)+C_{24}^4H_4(y_2)H_4(y_4)+C_{13}^3C_{24}H_3(y_1)H_1(y_2)H_3(y_3)H_1(y_4)\\
&+ C_{13}^2C_{24}^2H_2(y_1)H_2(y_2)H_2(y_3)H_2(y_4) + C_{13}C_{24}^3H_1(y_1)H_3(y_2)H_1(y_3)H_3(y_4)]\\
&+ \tfrac{1}{5!}[C_{13}^5H_5(y_1)H_5(y_3) + C_{24}^5H_5(y_2)H_5(y_4) + C_{13}^4C_{24}H_4(y_1)H_1(y_2)H_4(y_3)H_1(y_4)\\
&+ C_{13}^3C_{24}^2H_3(y_1)H_2(y_2)H_3(y_3)H_2(y_4) + C_{13}^2C_{24}^3H_2(y_1)H_3(y_2)H_2(y_3)H_3(y_4)\\
&+ C_{13}C_{24}^4H_1(y_1)H_4(y_2)H_1(y_3)H_4(y_4)] + \tfrac{1}{6!}[C_{13}^6H_6(y_1)H_6(y_3) + C_{24}^6H_6(y_2)H_6(y_4)\\
&+ C_{13}^5C_{24}H_5(y_1)H_1(y_2)H_5(y_3)H_1(y_4) + C_{13}^4C_{24}^2H_4(y_1)H_2(y_2)H_4(y_3)H_2(y_4)
\end{aligned}$$

$$+ C_{13}^{3}C_{24}^{3}H_3(y_1)H_3(y_2)H_3(Y_3)H_3(y_4) + C_{13}^{2}C_{24}^{4}H_2(y_1)H_4(y_2)H_2(y_3)H_4(y_4)$$

$$+ C_{13}C_{24}^{5}H_1(y_1)H_5(y_2)H_1(y_3)H_5(y_4)] + \;.\;\;.\;\;.\} \; f(\underset{\sim}{y}\colon \underset{\sim}{0},I) \;.$$

Before integrating (3.3.30), notice that $\phi(x)$ is an even function, i.e., $\phi(x) = \phi(-x)$. Also, $H_n(x)$ is an even function when n is even, and is an odd function when n is odd, i.e.,

$$H_n(-x) = \begin{cases} H_n(x) & \text{, n even} \\ -H_n(x) & \text{, n odd} \end{cases} \;.$$

Since $a_1=-b_1$ and $a_3=-b_3$, we have from Lemma 9,

$$\int_{a_i}^{b_i} H_n(x)\phi(x)dx = H_{n-1}(a_i)\phi(a_i)-H_{n-1}(b_i)\phi(b_i) \tag{3.3.31}$$

$$= H_{n-1}(-b_i)\phi(-b_i)-H_{n-1}(b_i)\phi(b_i)$$

$$= \begin{cases} -2H_{n-1}(b_i)\phi(b_i) & \text{if n is even (n-1 odd)} \\ 0 & \text{if n is odd (n-1 even)} \end{cases} \;.$$

Now integrate (3.3.30) using (3.3.31), Lemma 9 and Lemma 13 to obtain the desired result. □

In Chapter 4, each of the approximate methods outlined in this section will be applied to models similar to (3.3.4), and numerical powers of the two tests compared. Also, a study designed in part to determine the accuracy of these approximations is conducted. This is accomplished by comparing the approximate power of the test to the power generated from a Monte-Carlo simulation.

CHAPTER IV

A POWER COMPARISON OF THE BOX-PIERCE "PORTMANTEAU", MAX-χ^2, AND d TESTS

In Sections 3.1 and 3.2 we presented two testing procedures, the Box-Pierce "portmanteau" and max-χ^2 tests, designed to detect serial correlation of the errors in the presence of a general alternative error structure. In Chapter 4, an extensive power simulation study is undertaken. Our goal is three-fold. First, we wish to substantiate the conjecture that the Box-Pierce and max-χ^2 tests attain high power in the general case. Secondly, we investigate the specific types of alternative error models for which one test outperforms the other. Finally, the accuracy of the asymptotic power approximations (Section 3.3) is determined by comparing the simulated power to the approximate power.

4.1 Monte Carlo Simulations

Recall the stationary AR model of order p with lag coefficients $(\emptyset_1, \emptyset_2, \ldots, \emptyset_p)$ given by equation (3.3.1). In order to obtain the simulated powers of the tests via the Monte Carlo method, a technique for generating the sequence of observations $Z_1, Z_2, \ldots, Z_n$ needs to be developed. We adopt a procedure in which the first p-variates $Z_1, Z_2, \ldots, Z_p$ are generated so that they possess the desired covariance structure of the general alternative error model under consideration. The remaining n-p variates $Z_{p+1}, Z_{p+2}, \ldots, Z_n$ are then calculated recursively from equation (3.3.1). However, the formation of the covariance structure first entails computing the vector of true autocorrelations,

$\underset{\sim}{\rho}$. Given the lag coefficients $(\phi_1, \phi_2, \ldots, \phi_p)$ of the model, the vector $\underset{\sim}{\rho}$ is easily obtained. The relationship between $\underset{\sim}{\rho}$ and the lag coefficients is given in the following lemma.

Lemma 14: Consider the stationary p^{th} order AR model

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t, \qquad (4.1.1)$$

where $\{\varepsilon_t\}$ is an uncorrelated series with mean 0 and variance σ^2, and the vector of lag coefficients $\underset{\sim}{\phi}' = (\phi_1, \phi_2, \ldots, \phi_p)$ is known. Let $\underset{\sim}{\gamma}$ and $\underset{\sim}{\rho}$ represent the vectors of true autocovariances and true autocorrelations, respectively, where $\gamma_\nu = E\{Z_t Z_{t+\nu}\}$, $\rho_\nu = \gamma_o^{-1}\gamma_\nu$, $\nu = 0, 1, 2, \ldots$. Then $A\underset{\sim}{\rho} = \underset{\sim}{\phi}$, where the elements of the matrix A are given by:

$$a_{ii} = \begin{cases} 1 - \phi_{2i} & \text{if } 2i \leq p \\ 1 & \text{if } 2i > p \end{cases}$$

$$a_{ij} = \begin{cases} -\phi_{i+j} & \text{if } i+j \leq p,\ i < j \\ 0 & \text{if } i+j > p,\ i < j \end{cases}$$

$$a_{ij} = \begin{cases} -\phi_{i+j} - \phi_{i-j} & \text{if } i+j \leq p,\ i > j \\ -\phi_{i-j} & \text{if } i+j > p,\ i > j \end{cases}$$

Proof: Multiply Z_t in model (4.1.1) by $Z_{t+\nu}$, $\nu = 1, 2, 3, \ldots, p$, and take expectations to obtain the true autocovariances:

$$\begin{aligned}
\gamma_1 &= \phi_1\gamma_o + \phi_2\gamma_1 + \phi_3\gamma_2 + \phi_4\gamma_3 + \phi_5\gamma_4 + \ldots + \phi_p\gamma_{p-1} \\
\gamma_2 &= \phi_1\gamma_1 + \phi_2\gamma_o + \phi_3\gamma_1 + \phi_4\gamma_2 + \phi_5\gamma_3 + \ldots + \phi_p\gamma_{p-2} \\
\gamma_3 &= \phi_1\gamma_2 + \phi_2\gamma_1 + \phi_3\gamma_o + \phi_4\gamma_1 + \phi_5\gamma_2 + \ldots + \phi_p\gamma_{p-3} \\
&\ \vdots \\
\gamma_p &= \phi_1\gamma_{p-1} + \phi_2\gamma_{p-2} + \phi_3\gamma_{p-3} + \phi_4\gamma_{p-4} + \phi_5\gamma_{p-5} + \ldots + \phi_p\gamma_o.
\end{aligned}$$

Now utilize $\rho_\nu = \gamma_o^{-1}\gamma_\nu$ to obtain expressions for the true autocorrelations:

$$\rho_1 = \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + \phi_4\rho_3 + \phi_5\rho_4 + \cdots + \phi_p\rho_{p-1}$$

$$\rho_2 = \phi_1\rho_1 + \phi_2 + \phi_3\rho_1 + \phi_4\rho_2 + \phi_5\rho_3 + \cdots + \phi_p\rho_{p-2}$$

$$\vdots$$

$$\rho_p = \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \phi_3\rho_{p-3} + \phi_4\rho_{p-4} + \phi_5\rho_{p-5} + \cdots + \phi_p \; .$$

Solving the above equations for the lag parameters $\phi_1, \phi_2, \phi_3, \ldots \phi_p$ gives:

$$\phi_1 = (1-\phi_2)\rho_1 - \phi_3\rho_2 - \phi_4\rho_3 - \phi_5\rho_4 - \cdots - \phi_p\rho_{p-1}$$

$$\phi_2 = -(\phi_1+\phi_3)\rho_1 + (1-\phi_4)\rho_2 - \phi_5\rho_3 - \cdots - \phi_p\rho_{p-2}$$

$$\vdots$$

$$\phi_p = -\phi_{p-1}\rho_1 - \phi_{p-2}\rho_2 - \phi_{p-3}\rho_3 \cdots + \rho_p \quad ,$$

or in matrix form:

$$\underset{\sim}{\phi} = \begin{bmatrix} (1-\phi_2) & -\phi_3 & -\phi_4 & \cdots & -\phi_{p-1} & \phi_p & 0 \\ -(\phi_1+\phi_3) & (1-\phi_4) & -\phi_5 & \cdots & -\phi_p & 0 & 0 \\ \vdots & & & & & & \\ -\phi_{p-1} & -\phi_{p-2} & -\phi_{p-3} & \cdots & -\phi_2 & -\phi_1 & 1 \end{bmatrix} \underset{\sim}{\rho} \; .$$

To see this more clearly, consider the case, p=4. Expressions for the first four true autocorrelations are given by:

$$\rho_1 = \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + \phi_4\rho_3$$

$$\rho_2 = \phi\rho_1 + \phi_2 + \phi_3\rho_1 + \phi_4\rho_2$$

$$\rho_3 = \phi_1\rho_2 + \phi_2\rho_1 + \phi_3 + \phi_4\rho_1$$

$$\rho_4 = \phi_1\rho_3 + \phi_2\rho_2 + \phi_3\rho_1 + \phi_4$$

Now solve each of the above equations, respectively, for ϕ_j, j = 1, 2, 3, 4 to obtain:

$$\begin{aligned}
\phi_1 &= (1-\phi_2)\rho_1 - \phi_3\rho_2 - \phi_4\rho_3 \\
\phi_2 &= -(\phi_1+\phi_3)\rho_1 + (1-\phi_4)\rho_2 \\
\phi_3 &= -(\phi_2+\phi_4)\rho_1 - \phi_1\rho_2 + \rho_3 \\
\phi_4 &= -\phi_3\rho_1 - \phi_2\rho_2 - \phi_1\rho_3 + \rho_4.
\end{aligned}$$

In matrix form, we have:

$$\underset{\sim}{\phi} = \begin{bmatrix} 1-\phi_2 & -\phi_3 & -\phi_4 & 0 \\ -(\phi_1+\phi_3) & 1-\phi_4 & 0 & 0 \\ -(\phi_2+\phi_4) & -\phi_1 & 1 & 0 \\ -\phi_3 & -\phi_2 & -\phi_1 & 1 \end{bmatrix} \underset{\sim}{\rho} \quad ,$$

which agrees with the lemma result. □

The simulation technique used to generate the powers of the tests in Sections 4.1.1 and 4.1.2 is outlined in Table 4.1.

In Section 4.1.1 we examine the powers of the Box-Pierce and max-χ^2 tests when the vector of sample autocorrelations $\underset{\sim}{r}$ is known, i.e., when the random errors $\{Z_t\}$ are observable. Of course, in practice the $\{Z_t\}$ are unobservable. We have indicated (Section 3.3.2) that the researcher must instead calculate the vector of estimated sample autocorrelations, $\hat{\underset{\sim}{r}}$, using the residuals $\{\hat{Z}_t\}$ computed from least squares regression. Power comparisons of the Box-Pierce and max-χ^2 test in this more practical case are discussed in Section 4.1.2. These power comparisons also include the Durbin-Watson d test, which was designed to test regression residuals.

4.1.1 Observable Residuals: No Regression

In this section we present a power simulation study assuming that the residuals $\{Z_t\}$ from model (4.1.1) are observable. The objective is to detect power trends which may extend to the more practical instance in which powers are calculated using the residuals $\{\hat{Z}_t\}$ computed from least squares regression.

TABLE 4.1

MONTE CARLO SIMULATION

ERROR MODEL: $Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \ldots + \emptyset_p Z_{t-p} + \varepsilon_t$,

where $\{\varepsilon_t\}$ is White Noise

STEP 1: Utilizing Lemma 14, compute the vector of true autocorrelations

$$\underset{\sim}{\rho} = A^{-1} \underset{\sim}{\emptyset} ,$$

where $\underset{\sim}{\rho} = (\rho_1, \rho_2, \ldots, \rho_p)'$, $\underset{\sim}{\emptyset} = (\emptyset_1, \emptyset_2, \ldots, \emptyset_p)'$ and the elements of A are given by:

$$a_{ii} = \begin{cases} 1-\emptyset_{2i} & \text{if } 2i \leq p \\ 1 & \text{if } 2i > p \end{cases}$$

$$a_{ij} = \begin{cases} -\emptyset_{i+j} & \text{if } i + j \leq p,\ i<j \\ 0 & \text{if } i + j > p,\ i<j \\ -\emptyset_{i+j}-\emptyset_{i-j} & \text{if } i + j \leq p,\ i>j \\ -\emptyset_{i-j} & \text{if } i + j > p,\ i>j \end{cases} .$$

STEP 2: Generate the covariance structure, Γ, of the first p-variates $(Z_1, Z_2, \ldots, Z_p)$ as follows:

$$\Gamma = \begin{bmatrix} \gamma_o & \gamma_1 & \gamma_2 & \cdots & \gamma_{p-1} \\ \gamma_1 & \gamma_o & \gamma_1 & \cdots & \gamma_{p-2} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_o \end{bmatrix} ,$$

where $\gamma_\nu = E\{Z_t \cdot Z_{t+\nu}\} = \gamma_o \cdot \rho_\nu$, $\nu = 1, 2, 3, \ldots, p$ and

TABLE 4.1--Continued

$$\gamma_o = \phi_1\gamma_1 + \phi_2\gamma_2 + \ . \ . \ . + \phi_p\gamma_p + \sigma^2 = (\phi_1\rho_1 + \phi_2\rho_2 + \ . \ . \ . + \phi_p\rho_p)\gamma_o + \sigma^2$$

$$= \frac{\sigma^2}{1-(\phi_1\rho_1 + \phi_2\rho_2 + \ . \ . \ . + \phi_p\rho_p)} \ .$$

Without loss of generality, take $E(\varepsilon_t^2) = \sigma^2 = 1$.

STEP 3: Write $\Gamma = TT'$ where T is a pxp lower triangular matrix. Solve for the elements of T using a computer routine.

STEP 4: Generate the first p random variates $Z_1, Z_2, . . ., Z_p$ by computing $\underset{\sim}{Z}_p = T \cdot \varepsilon_p$, where $\underset{\sim}{Z}_p = (Z_1, Z_2, . . ., Z_p)'$, and where $\underset{\sim}{\varepsilon}_p = (\varepsilon_1, \varepsilon_2, . . ., \varepsilon_p)'$ is a vector of independent normal random variables with 0 and variance 1. (Many computer packages have routines which generate N(0,1) random variates.) Note that $\mathrm{Cov}\{\underset{\sim}{Z}_p\} = E\{\underset{\sim}{Z}_p\underset{\sim}{Z}_p'\} = TE\{\underset{\sim}{\varepsilon}_p\underset{\sim}{\varepsilon}_p'\}T' = TT' = \Gamma$, and thus the first p random variates generated have the desired covariance structure.

STEP 5: Compute the remaining n-p variates, $Z_{p+1}, Z_{p+2}, . . ., Z_n$, from the recursive relations given by model (4.1.1), i.e.,

$$Z_{p+1} = \phi_1 Z_p + \phi_2 Z_{p-1} + \ . \ . \ . + \phi_p Z_1 + \varepsilon_{p+1}$$

$$Z_{p+2} = \phi_1 Z_{p+1} + \phi_2 Z_p + \ . \ . \ . + \phi_p Z_2 + \varepsilon_{p+2}$$

$$\vdots$$

$$Z_n = \phi_1 Z_{n-1} + \phi_2 Z_{n-2} + \ . \ . \ . + \phi_p Z_{n-p} + \varepsilon_n \ ,$$

where $\varepsilon_{p+1}, . . ., \varepsilon_n$ are distributed multi-variate normal with mean 0 and covariance I.

STEP 6: Use the generated random errors $Z_1, Z_2, . . ., Z_n$ to calculate the first K sample autocorrelations

TABLE 4.1--Continued

$$r_1, r_2, \ldots, r_K, \text{ where } r_j = \frac{\sum_{t=j+1}^{n} Z_t Z_{t-j}}{\sum_{t=1}^{n} Z_t} , \quad 1 \leq j \leq K.$$

STEP 7: Compute the test statistics, $T^{(BP)} = n \sum_{j=1}^{K} r_j$, and $T^{(m)} = n \max_{1 \leq j \leq K} r_j$

Determine whether the calculated values of the test statistics fall in their respective rejection regions:

$$T(BP) > \chi^2_{K,\alpha}$$

$$T^{(m)} > \left[\Phi^{-1}\left\{\frac{1}{2} + \frac{(1-\alpha)^{1/K}}{2}\right\}\right]^2 .$$

STEP 8: Repeat steps 1 - 7 N times where we selected N = 1000.

STEP 9: Compute the simulated power for each test as follows:

$$\text{Power} = \frac{\text{number of rejections}}{1000} .$$

A list of the controllable parameters in the study is given below:

n = number of residuals generated

α = probabilility of a Type I error

K = number of sample autocorrelations included in the Box-Pierce and max- χ^2test statistics

p = order of the autoregressive model

$(\phi_1, \phi_2, \ldots, \phi_p)$ = values of the lag coefficients in model (4.1.1)

m = number of nonzero lag coefficients in model (4.1.1)

$(j_1, j_2, \ldots, j_m)$ = lags associated with nonzero lag coefficients.

Of course, there exists an infinite number of parameter combinations. For our study, we have attempted to select those model parameters which are intuitively appealing to the time-series analyst and/or econometrician, i.e., those which also have practical applications. Our basic error model takes the form of equation (3.3.4). The two versions of this model used in the study are: (1) the first order AR process with lag J

$$Z_t = \phi Z_{t-J} + \varepsilon_t , \qquad (4.1.2)$$

and (2) the second order AR process with lags, J_1 and J_2

$$Z_t = \phi_{J_1} Z_{t-J_1} + \phi_{J_2} Z_{t-J_2} + \varepsilon_t .$$

The parameter values of each model are given in Table 4.2 and Table 4.3, respectively. In order to present the results of the study in a clear fashion, we do not give tables of the numerical power simulations in this section. Instead, we sketch power curves for various cases which are typical of those considered and which also best emphasize the main results. Tables of all numerical results obtained in the study are included in the Appendix which follows the final chapter. However, when summarizing the simulation results, we will at times refer the reader to these tables.

TABLE 4.2

SIMULATION STUDY
ERROR MODEL 1: $Z_t = \emptyset Z_{t-J} + \varepsilon_t$

Parameters:

(J,K)	Ø	n	α
(2,4)	-.9	50	.05
(3,4)	-.7	200	
(4,4)	-.5		
(5,4)*	-.3		
(2,12)	-.1		
(3,12)	0		
(4,12)	.1		
(12,12)	.3		
(12,24)	.5		
	.7		
	.9		

*NOTE: Powers are generated for all of the above combinations of {(J,K) Ø, n, α} with the excpetion of the case {(5,4), Ø, 200, .05}.

TABLE 4.3

SIMULATION STUDY

ERROR MODEL 2: $Z_t = \phi_{J_1} Z_{t-J_1} + \phi_{J_2} Z_{t-J_2} + \varepsilon_t$

Parameters:	$(J_1,J_2;K)$	ϕ_{J_1},ϕ_{J_2}	n	α
	(1,2;2)	(-.5,.3)	50	.05
	(1,2;12)	(.5, .3)	200	
		(-1.0,-.6)		
		(1.0,-.6)		
		(.1,-.7)		
		(.1,.7)		
		(-.3,.1)		
		(.7,.3)		
		(-.7,.3)		
	(1,3;4)	(.1,-.9)	50	.05
		(.1,.7)	200	
		(.5,-1.2)		
		(-.5,-.3)		
		(1.0,.1)		
		(-1.0,.1)		
		(.5,-.3)		
		(-1.0,.3)		
		(.3,-.1)		

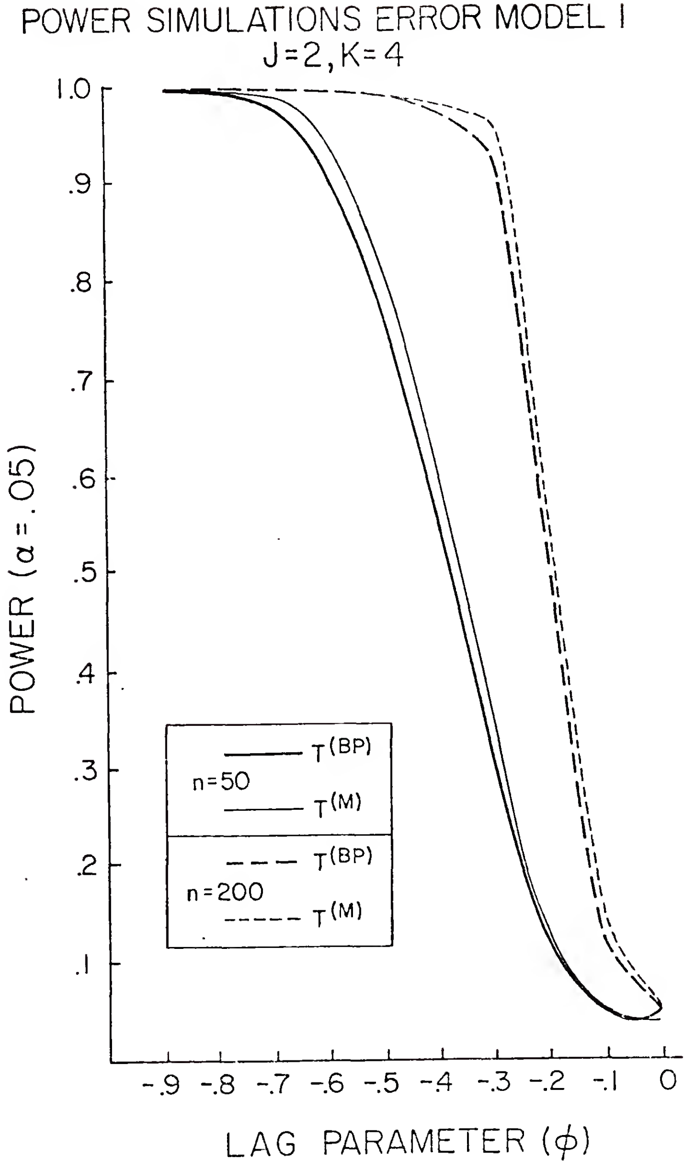

POWER SIMULATIONS ERROR MODEL I
J=2,K=4
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
POWER (α = .05)
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (φ)
n=50
T(BP)
T(M)
n=200
T(BP)
T(M)

FIGURE 4.1a

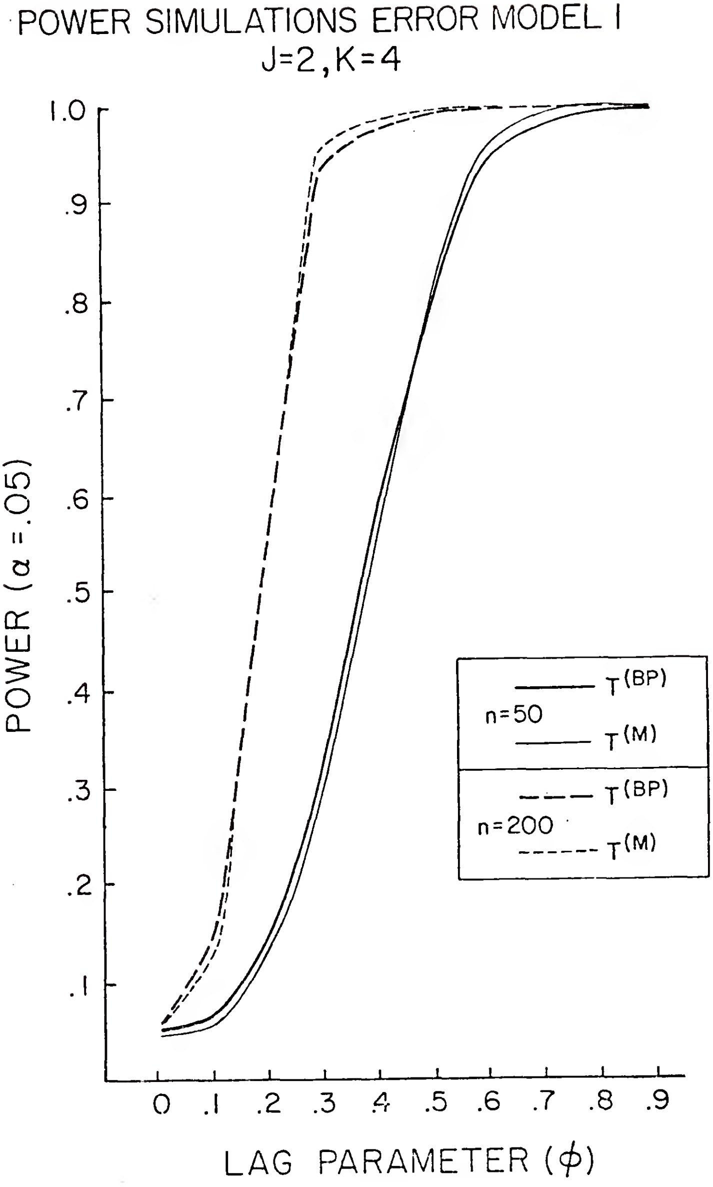

POWER SIMULATIONS ERROR MODEL I
J=2, K=4
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (ϕ)
n=50
T(BP)
T(M)
n=200
T(BP)
T(M)

FIGURE 4.1b

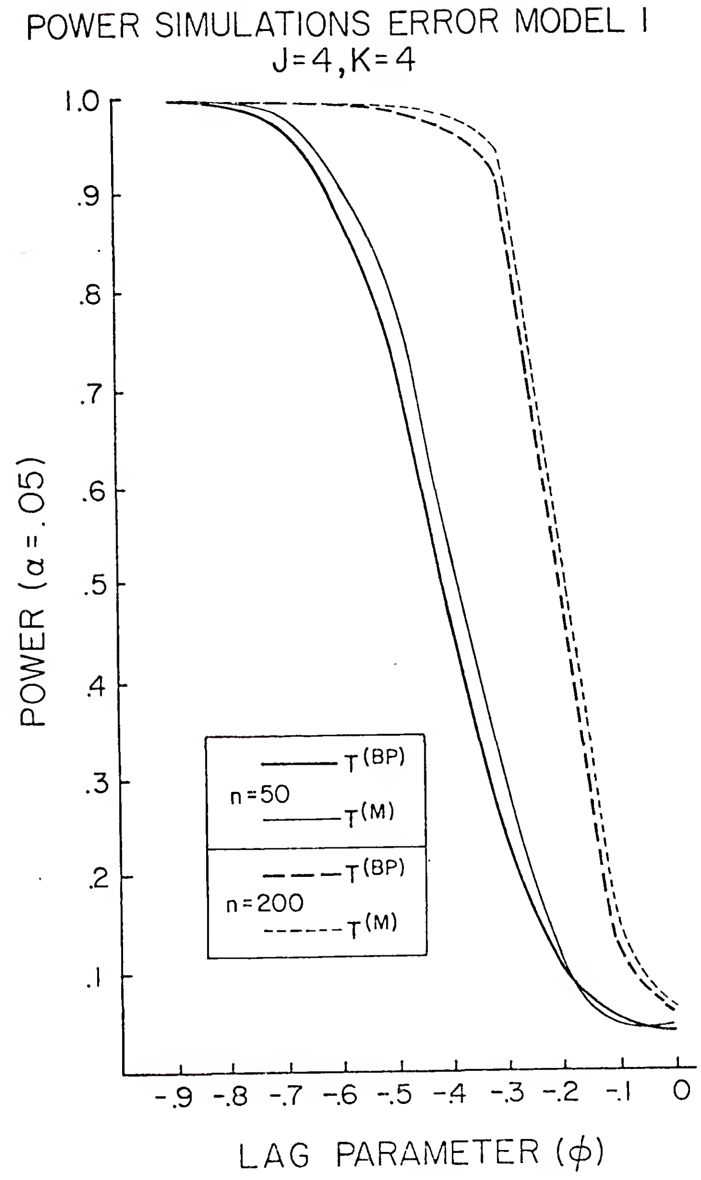

POWER SIMULATIONS ERROR MODEL I
J=4, K=4
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
POWER (α = .05)
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (ϕ)
n=50
$T^{(BP)}$
$T^{(M)}$
n=200
$T^{(BP)}$
$T^{(M)}$

FIGURE 4.2a

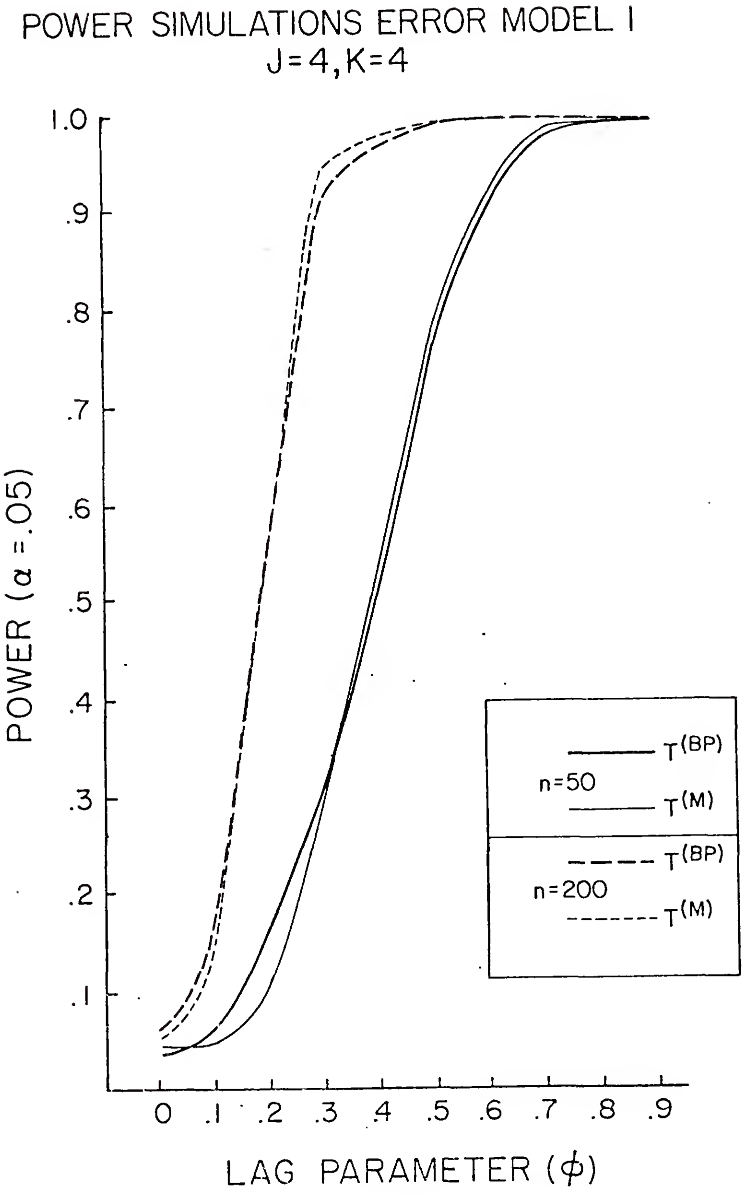
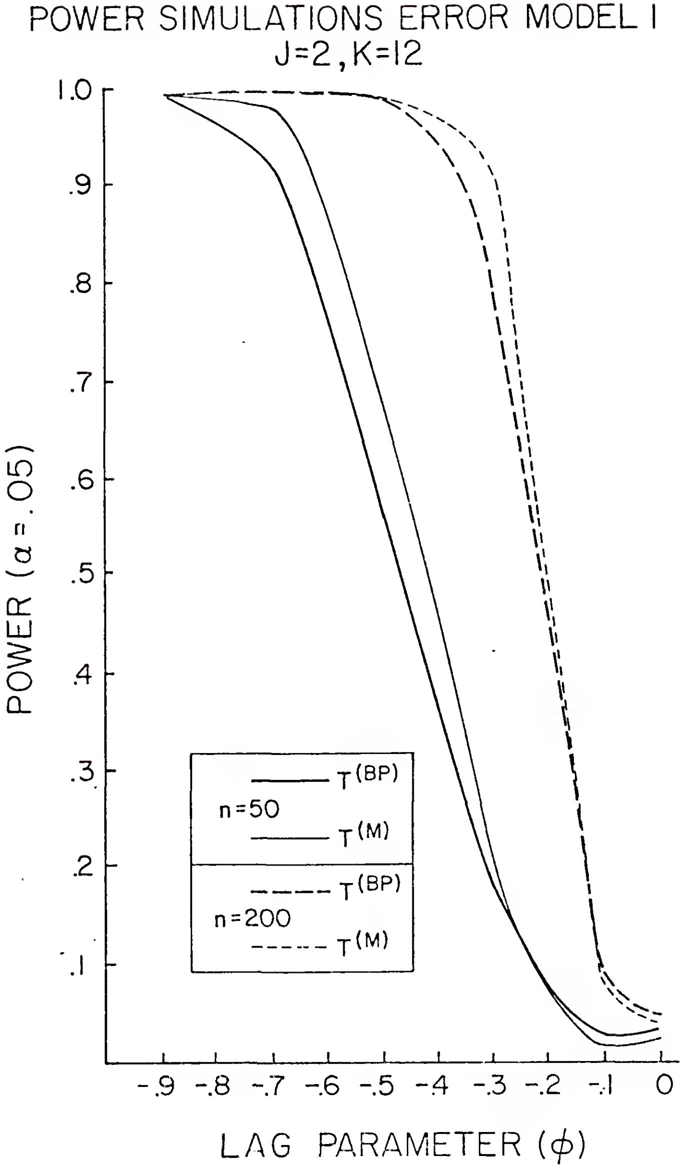

POWER SIMULATIONS ERROR MODEL I
J=4,K=4
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (φ)
n=50
T(BP)
T(M)
n=200
T(BP)
T(M)

FIGURE 4.2b



POWER SIMULATIONS ERROR MODEL I
J=2,K=12

FIGURE 4.3a

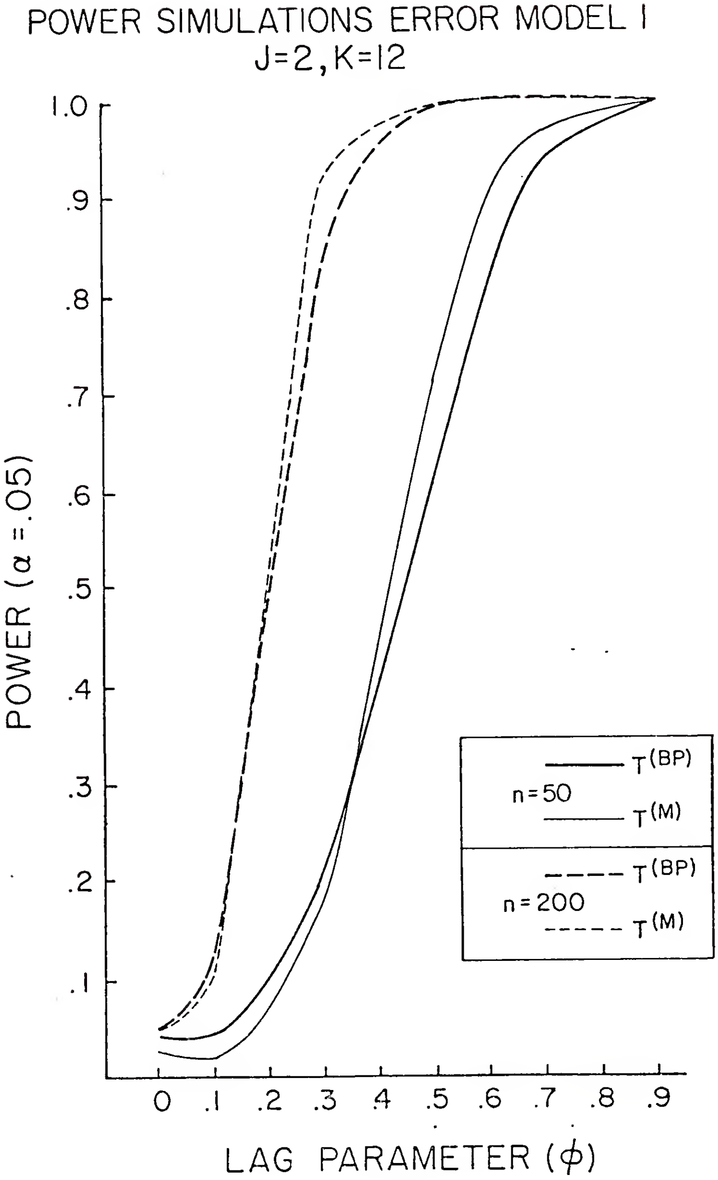

POWER SIMULATIONS ERROR MODEL I
J=2, K=12
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (ϕ)
n=50
T(BP)
T(M)
n=200
T(BP)
T(M)

FIGURE 4.3b

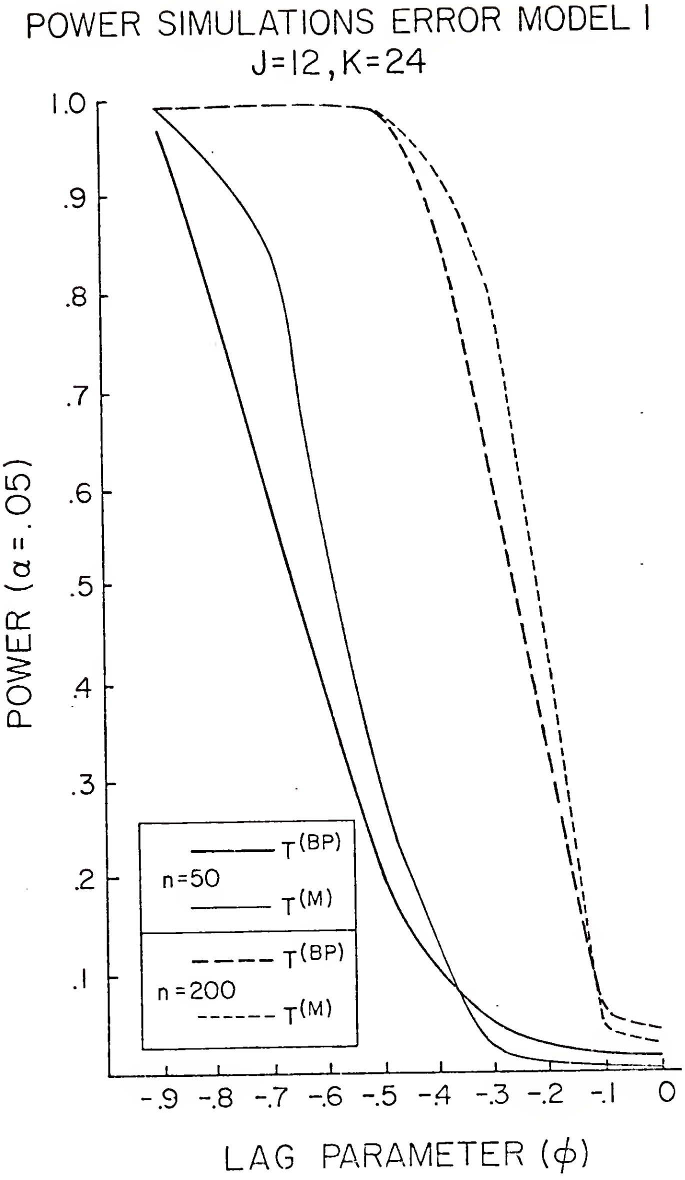

POWER SIMULATIONS ERROR MODEL I
J=12, K=24
POWER (α=.05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (φ)
n=50
T(BP)
T(M)
n=200
T(BP)
T(M)

FIGURE 4.4a

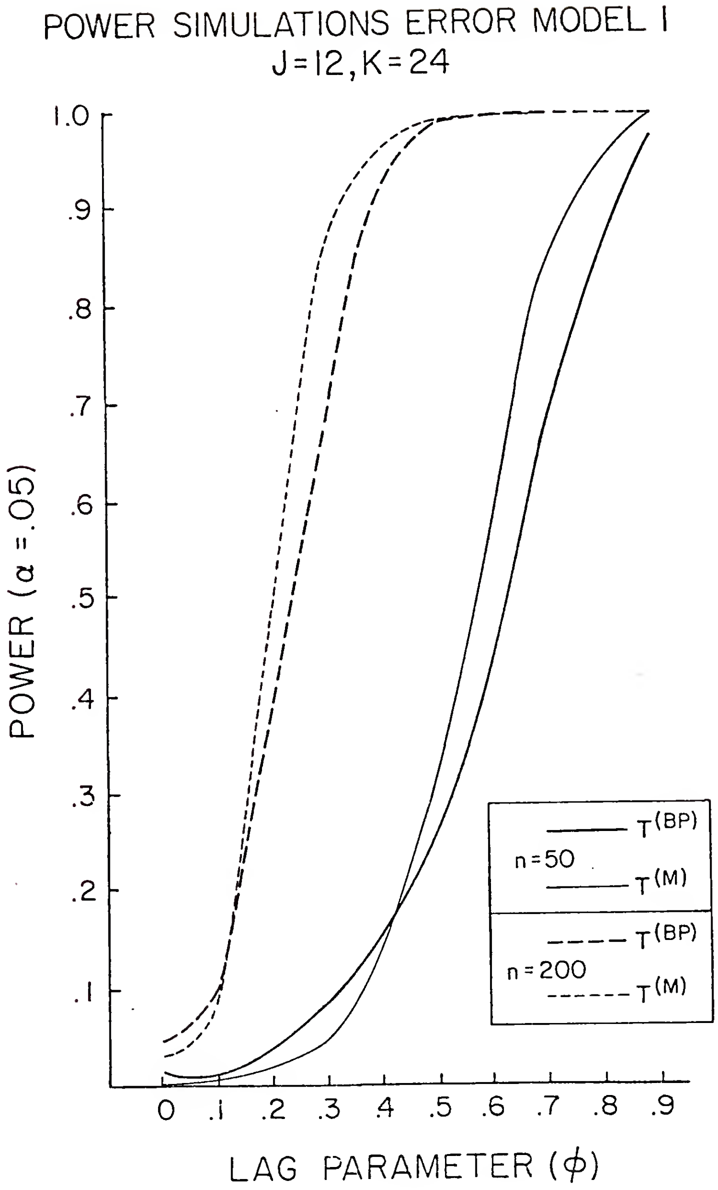

POWER SIMULATIONS ERROR MODEL I
J=12, K=24
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (ϕ)
n=50
$T^{(BP)}$
$T^{(M)}$
n=200
$T^{(BP)}$
$T^{(M)}$

FIGURE 4.4b

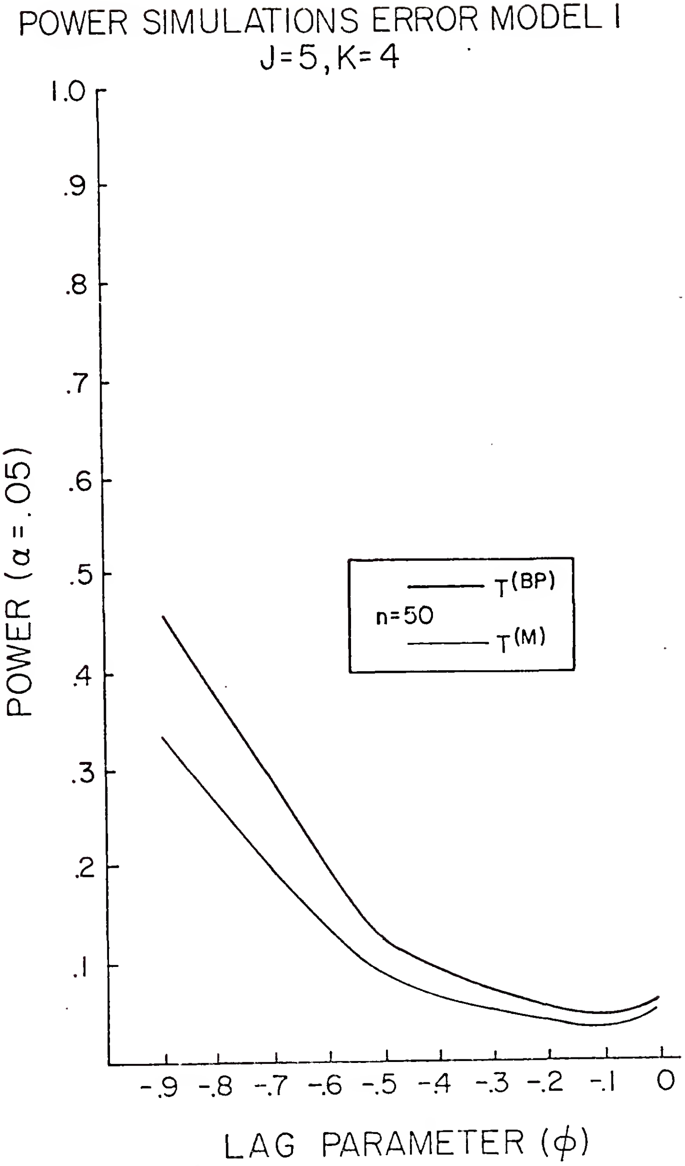

POWER SIMULATIONS ERROR MODEL I
J=5, K=4
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
POWER (α = .05)
$T^{(BP)}$
n=50
$T^{(M)}$
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (φ)

FIGURE 4.5a

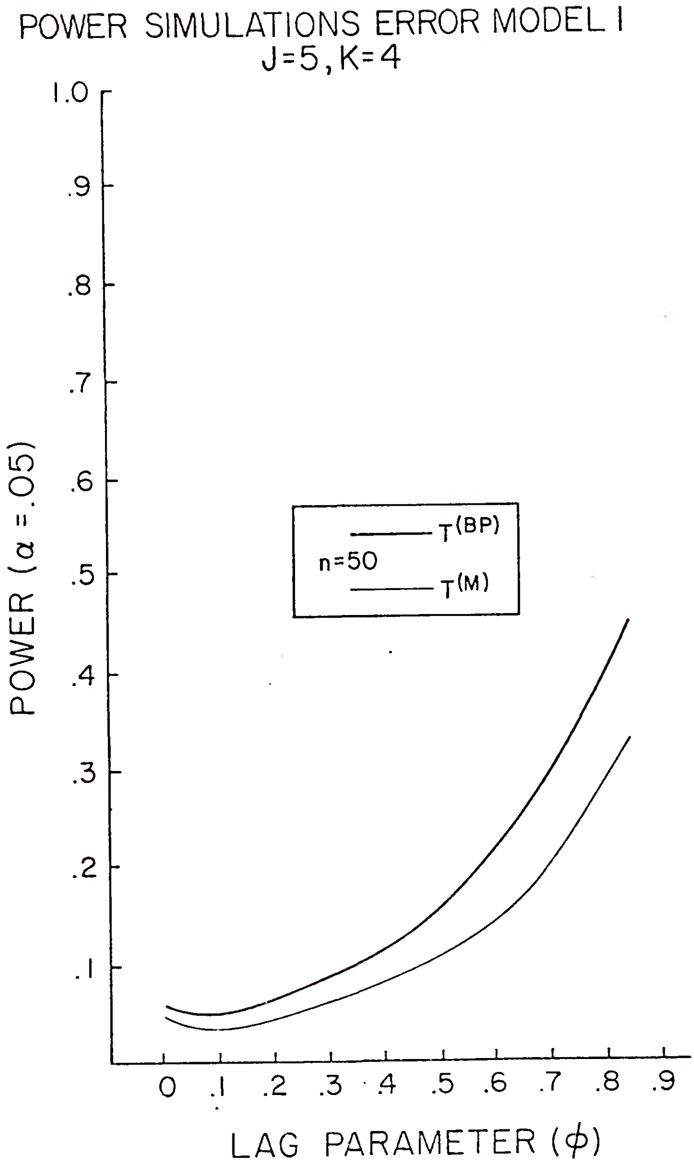

POWER SIMULATIONS ERROR MODEL I
J=5, K=4
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
POWER (α = .05)
$T^{(BP)}$
n=50
$T^{(M)}$
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (ϕ)

FIGURE 4.5b

We first discuss the power simulation results for error model 1 (see Tables A-1 - A-9 in Appendix). Note that for n=50, Figures 4.1 - 4.4 show that moderate and large values of the lag parameter (in absolute value), i.e., $.5 \leq |\emptyset| \leq 1$, lead to a high rate of rejection of the null hypothesis of white noise for both the Box-Pierce and max-χ^2 tests. When the sample size is increased to 200, high rejection rates result at values of the lag parameter .3 or larger (in absolute value). Except for the case J=5, K=4 (Figure 4.5), these results hold for all other cases considered. For these models, then, this simulation study clearly supports our supposition that both tests attain high power under a general alternative. Compare now Figure 4.1 and 4.3, i.e., the cases J=2, K=4 and J=2, K=12. Notice that the powers of the tests generally decrease as K, the number of autocorrelations included in the test statistic is increased from 4 to 12. And in all other cases considered, for a fixed lag J, the powers of the tests decreased as K increased. In contrast, if the researcher unknowingly selects K less than the smallest nonzero lag in the model (J in our model), the powers of the tests are reduced dramatically, as demonstrated in Figure 4.5 for the case J=5, K=4. For moderate and small values of the lag parameter, the powers of the tests are less than .10. Thus, the test user's choice of K is a delicate one. The researcher who conservatively selects a K much larger than the largest nonzero lag in the general alternative error model will fail to reject the null hypothesis of white noise when serial correlation exists at a more frequent rate. Alternatively, if K is chosen smaller than the lowest lag in the hypothesized error model, the tests have very low power.

Comparing the powers of the tests, Figures 4.1a - 4.4a reveal that the power of the max-χ^2 procedure dominates the power of the Box-Pierce test for almost all negative values of the lag parameter, or more specifically, for $\emptyset \leq .3$. Consider the case J=12, K=24 (Figure 4.4). Referring to Table A-8 in the Appendix, at $\emptyset = -.7$ and n=50, the max-χ^2 test with a simulated power of .853 greatly outperforms the Box-Pierce test (simulated power of .583). When analyzing yearly data, of course, it is very possible that the regression errors $\{Z_t\}$ follow an autoregressive scheme with nonzero lag 12. Thus, the practical significance of the result is clear. For small values of the lag parameter, i.e., $-.3 \leq \emptyset \leq .3$, the figures show that the Box-Pierce power is generally larger than the max-χ^2 power. However, this result seems much weaker than that previously stated since the powers attained by both tests for $\emptyset$ in this range are very low, and in some cases, near zero. The apparent trend in the powers of the tests for moderate and large values of the lag parameter at n=50 can be stated as follows: (a) for $.3 \leq \emptyset \leq .5$, the powers of the tests are very nearly equivalent; (b) for $\emptyset \geq .5$, the power of the max-χ^2 test is slightly higher than the Box-Pierce power. The max-χ^2 procedure also seems to be less sensitive to an increase in K. For example, compare Figures 4.1 and 4.3, the cases J=2, K=4 and J=2, K=12. At $\emptyset = .3$ and n=200 the Box-Pierce power drops from .947 (Table A-1) to .837 (Table A-4), while the drop in the power of the max-χ^2 test is only from .959 to .925.

Figure 4.5 represents the only case (J=5, K=4) in which the Box-Pierce test clearly outperforms the max-χ^2 test. In this instance, we expect both tests to perform poorly. With J=5, the first four true autocorrelations ρ_1, ρ_2, ρ_3, ρ_4 are zero, and thus each of the first four sample autocorrelations r_1, r_2, r_3, r_4 estimate a quantity known to be zero. Hence, both the sum and the maximum of the first four sample autocorrelations will be small and will rarely lead to a significant result. The only reason that the two tests obtain any power at all (i.e., power greater than .05) is due to v_{ii}, the variance of the asymptotic distribution of $\sqrt{n}r_i$, i = 1, 2, 3, 4. From Section 3.3.3, equation (3.3.9) gives $v_{ii} = \frac{1+\phi^2}{1-\phi^2}$, i = 1, 2, 3, 4, for the case J=5, K=4. As the magnitude of ϕ increases, v_{ii} also increases. Thus, even if $\rho_i = 0$, a large value of ϕ could possibly lead to large values of r_i due to random variation, and hence, a significant result. This unusual phenomenon will, of course, occur more often when using the Box-Pierce test, since the variability of the sum of the first four sample autocorrelations will be larger than the variability of the maximum. However, it should again be noted that the Box-Pierce test, while outperforming the max-χ^2 test, attains a very low and unappealing power.

A final point to be made is that for all cases considered in this part of the study, the powers of the tests appear to be very low (near 0) for values of the lag parameter near 0. However, one must remember that the critical regions used in this study were those asymptotic regions discussed in Section 3.3. In theory, when $\phi = 0$, the power of the test and significance level (α) of the test are identical. But a study of the n=50 results for $\phi = 0$ in Tables A-1 - A-9 of the Appendix reveals that the majority of the simulated powers are smaller than .05. Thus, the actual level of the test at n=50 is less than the asymptotic level (.05) used in the study, which in turn makes the power small for ϕ near 0. Notice that when n is increased to 200, powers at $\phi = 0$ are very close to .05, which indicates that the asymptotic level of the test statistics apparently have been attained.

Summarizing the main results for model 1:

(i) For moderate to large negative values of the lag parameter, the power of the max-χ^2 test dominates the power of the Box-Pierce test.

(ii) For $-.1 \leq \emptyset \leq .1$, the tests have very low power.

(iii) In most cases, when $\emptyset \geq .3$, the max-χ^2 test outperforms the Box-Pierce test.

(iv) The max-χ^2 test appears to be less sensitive (in terms of power) to changes in the number of sample autocorrelations included in the test statistic.

(v) When the number of sample autocorrelations K is chosen less than the smallest nonzero lag J both tests attain a very low power.

These results are not surprising, for model 1 is the type of alternative (one with few nonzero lag parameters) under which we hypothesized that the Box-Pierce procedure would frequently fail to reject the null hypothesis of white noise (see Section 3.2).

A close examination of the simulated powers for model 2 reveals no discernable trend. Thus, we omit power curves for these cases. For versions of model 2 considered, the observed difference in the powers of the two tests never exceeds .042 (see Tables A-10 - A-12 of the Appendix). Although it appears that the Box-Pierce test and max-χ^2 test attain approximately the same power for these second order models, one should not assume that the powers will be identical for all second order models. A more extensive simulation study needs to be undertaken.

As previously stated, the regression errors $\{Z_t\}$ in any practical case will be unobservable. In the following section, we examine the simulated powers of the two tests when the test statistics are computed using residuals calculated from least squares regression.

4.1.2 Estimated Residuals After Regression

Recall the problem (stated in Section 1.1) of incorrectly applying least squares inferential methodology to the parameters of the regression model:

$$\underset{\sim}{Y}_{(nx1)} = X_{(nxg)} \cdot \underset{\sim}{\beta}_{(gx1)} + \underset{\sim}{Z}_{(nx1)} \quad , \tag{4.1.3}$$

where $Z_1, Z_2, \ldots, Z_n$ are unobservable random errors. In order to utilize the Box-Pierce, max-χ^2 , or Durbin-Watson d procedure as a test for independence of the $\{Z_t\}$, the researcher must first compute the vector of least squares residuals $\underset{\sim}{\hat{Z}} = (\hat{Z}_1, \hat{Z}_2, \ldots, \hat{Z}_n)$, where

$$\underset{\sim}{\hat{Z}} = \underset{\sim}{Y} - X \cdot \underset{\sim}{\hat{\beta}}$$

and

$$\underset{\sim}{\hat{\beta}} = (X'X)^{-1}X'\underset{\sim}{Y} \quad .$$

These residuals are then substituted for $\{Z_t\}$ in the formulas for the test statistics, i.e.,

$$T^{(BP)} = n \cdot \sum_{i=1}^{K} \hat{r}_j^{\,2} \quad ,$$

$$T^{(m)} = n \max_{1 \leq j \leq K} \hat{r}_j^{\,2} \quad ,$$

where

$$\hat{r}_j = \Big(\sum_{t=j+1}^{n} \hat{Z}_t \, \hat{Z}_{t-j}\Big) \Big/ \Big(\sum_{t=1}^{n} \hat{Z}_t^2\Big) \quad ,$$

and

$$d = \frac{\sum_{t=2}^{n} (\hat{Z}_t - \hat{Z}_{t-1})^2}{\sum_{t=1}^{n} \hat{Z}_t^2}$$

Our power simulations, in this regression case, now require additional steps to those given in Section 4.1, for there are two models to consider: the general regression model (4.1.3) of interest and the autoregressive pattern of the random errors under the alternative hypothesis of dependence (4.1.1). We outline the simulation technique in Table 4.4.

Four regression models are considered in this study.

MODEL A: $Y_t = \beta_o + \beta_1 t + Z_t, \quad t = 1, 2, 3, \ldots, n$

MODEL B: $Y_t = \beta_o + \beta_1 t + \beta_2 t^2 + Z_t, \quad t = 1, 2, 3, \ldots, n$

MODEL C: $Y_t = \beta_o + \beta_1 X_t + \beta_2 t + Z_t$

$X_t = \gamma_o + \gamma_1 X_{t-1} + \eta_t, \quad t = 1, 2, 3, \ldots, n$

where $\{\eta_t\} \sim$ iid $N(0,1)$ and $|\gamma_1| < 1$

MODEL D: $Y_t = \beta_o + \beta_1 X_{1t} + \beta_2 X_{2t} + Z_t$,

$X_{1t} = \alpha_o + \alpha_1 X_{1,t-1} + \omega_{1t}$,

$X_{2t} = \gamma_o + \gamma_1 X_{2,t-1} + \omega_{2t}$,

where $|\alpha_1| < 1$, $|\gamma_1| < 1$ and $\underset{\sim}{\omega} = (\omega_{1t}, \omega_{2t})' \sim N\{\underset{\sim}{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\}$

Powers of the Box-Pierce, max-χ^2, and Durbin-Watson d test are simulated for each regression model under the first order lag J AR alternative:

$$Z_t = \phi Z_{t-J} + \varepsilon_t \quad ,$$

that is, model 1 of Section 4.1.1. Also estimated is the probability that the d statistic falls in the "inconclusive" region (Section 1.2.1). Our format for presenting the simulation results is similar to that of the previous section. Tables of the parameter values for each regression model are followed by power curve figures which are typical of the cases considered. We again include all numerical power results in table form in the Appendix.

TABLE 4.4

MONTE CARLO SIMULATION

REGRESSION MODEL: $Y_t = \beta_o + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_g X_{gt} + Z_t$

ALTERNATIVE ERROR MODEL: $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t$, where $\{\varepsilon_t\}$ is white noise.

STEPS 1 - 5: Same as previously given in Table 4.1.

STEP 6: Use the generated random errors $\{Z_t\}$ from Step 5 to obtain simulated values of the observable random variables $\{Y_t\}$ from the regression model (4.1.1). (Assume $\beta_o, \beta_1, \ldots, \beta_g$ are known, and use these values to compute Y_t).

STEP 7: Apply the method of least squares to the regression model (4.1.1) in order to obtain $\hat{\underset{\sim}{\beta}}$.

STEP 8: Compute the vector of least squares residuals $\hat{\underset{\sim}{Z}}$.

STEP 9: Use the simulated estimates of the random errors to obtain estimates of the first K sample autocorrelations $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_K$.

STEP 10: Compute the three test statistics $T^{(BP)}, T^{(m)}$ and d. Determine whether the calculated values of the test statistics fall in their respective rejection regions:

$$T^{(BP)} > \chi^2_{K,\alpha}$$

$$T^{(m)} > \left[\Phi^{-1}\left\{\frac{1}{2} + \frac{(1-\alpha)^{1/K}}{2}\right\}\right]^2$$

$d < d^*_{L,\alpha/2}$ or $d > 4-d^*_{L,\alpha/2}$ (where $d^*_{L,\alpha/2}$ depends upon n-g and may be obtained from available D-W tables).

TABLE 4.4--Continued

STEP 11: Repeat steps 1 - 10 N times (we used N=1000 iterations).

STEP 12: Compute the simulated power for each test as follows:

$$\text{Power} = \frac{\text{number of rejections}}{1000} \quad .$$

Before presenting the results, we note the sources for some of the model parameters used in the study. The regression coefficients for model B were obtained from Durbin-Watson [10] , where the annual price of sugar Y_t was regressed over time. Regression model C is similar to the model studied by Fuller [13] , pp. 444, where

Y_t = the logarithm of employment (in thousands) of the cement industry,

X_t = the logarithm of the production index of the industry,

t = time in quarters.

A model, first discussed by Durbin-Watson [10] and later by Fuller [13], pp. 426, that attempts to explain the annual per capital consumption of spirits (Y_t) using the independent variables per capita income (X_{1t}) and price of spirits (X_{2t}) is similar to our regression model D. The inclusion of model D enables us to examine the behavior of the residuals when the independent variables (X_{1t}, X_{2t}) are highly correlated.

(The power simulation results for all cases considered in this section are given in Tables A-13 - A-28 of the Appendix.) The trend in the powers of the Box-Pierce and max-χ^2 test, as piectured in Figures 4.6 - 4.8, is almost identical to the non-regression case. Except for small positive values of the lag parameter $\emptyset$, the max-χ^2 power dominates the Box-Pierce power. Again, for fixed J, the performance of the max-χ^2 procedure appears to be less sensitive to an increase in K (see tables in the Appendix). The powers of the Box-Pierce and max-χ^2 test are, once more, near zero for small values of the lag parameter. At $n = 50$, the actual levels of the tests are less than the asymptotic level (.05).

TABLE 4.5

SIMULATION STUDY
REGRESSION MODEL A: $Y_t = \beta_o + \beta_1 t + Z_t$
ERROR MODEL 1: $Z_t = \emptyset Z_{t-J} + \varepsilon_t$

Parameters:	(J,K)	$\emptyset$	n	α	β_o	β_1
	(2,4)	-.9	50	.05	.4	4.4
	(3,4)	-.8				
	(4,4)	-.7				
	(1,12)	-.6				
	(4,12)	-.5				
		-.4				
		-.3				
		-.2				
		-.1				
		0				
		.1				
		.2				
		.3				
		.4				
		.5				
		.6				
		.7				
		.8				
		.9				

TABLE 4.6

SIMULATION STUDY

REGRESSION MODEL B: $Y_t = \beta_o + \beta_1 t + \beta_2 t^2 + Z_t$

ERROR MODEL 1: $Z_t = \emptyset Z_{t-J} + \varepsilon_t$

Parameters:	(J,K)	$\emptyset$	n	α	β_o	β_1	β_2
	(4,4)	-.9	50	.05	21.548	-.2544	.05727
	(1,12)	-.8					
	(4,12)	-.7					
		-.6					
		-.5					
		-.4					
		-.3					
		-.2					
		-.1					
		0					
		.1					
		.2					
		.3					
		.4					
		.5					
		.6					
		.7					
		.8					
		.9					

TABLE 4.7

SIMULATION STUDY

REGRESSION MODEL C: $Y_t = \beta_o + \beta_1 X_t + \beta_2 t + Z_t$

$X_t = \gamma_o + \gamma_1 X_{t-1} + \eta_t$, $\{\eta_t\}$ white noise

ERROR MODEL 1: $Z_t = \emptyset Z_{t-J} = \varepsilon_t$

Parameters:	(J,K)	Ø	n	α	β_o	β_1	β_2
	(2,4)	-.9	50	.05	.84	.098	-.00057
	(4,4)	-.8					
	(1,12)	-.7			γ_o	γ_1	
	(4,12)	-.6			.8	.8	
		-.5					
		-.4					
		-.3					
		-.2					
		-.1					
		0					
		.1					
		.2					
		.3					
		.4					
		.5					
		.6					
		.7					
		.8					
		.9					

TABLE 4.8

SIMULATION STUDY

REGRESSION MODEL D: $Y_t = \beta_o + \beta_1 X_{1t} + \beta_2 X_{2t} + Z_t$

$X_{1t} = \gamma_o + \gamma_1 X_{1,t-1} + \omega_{1t}$

$X_{2t} = \alpha_o + \alpha_1 X_{2,t-1} + \omega_{2t}$

where $(\omega_{1t},\omega_{2t})' \sim N\{0, \begin{bmatrix}1 & \rho \\ \rho & 1\end{bmatrix}\}$

ERROR MODEL 1: $Z_t = \emptyset Z_{t-J} + \varepsilon_t$

Parameters:

(J,K)	Ø	n	α	β_o	β_1	β_2
(2,4)	-.9	50	.05	2.14	.69	-.63
(4,4)	-.8					
(1,12)	-.7					
(4,12)	-.6					
	-.5					
	-.4					
	-.3					
	-.2					
	-.1					
	0					
	.1					
	.2					
	.3					
	.4					
	.5					
	.6					
	.7					
	.8					
	.9					

γ_o	γ_1	α_o	α_1	ρ
,8	.8	.3	-.5	.7

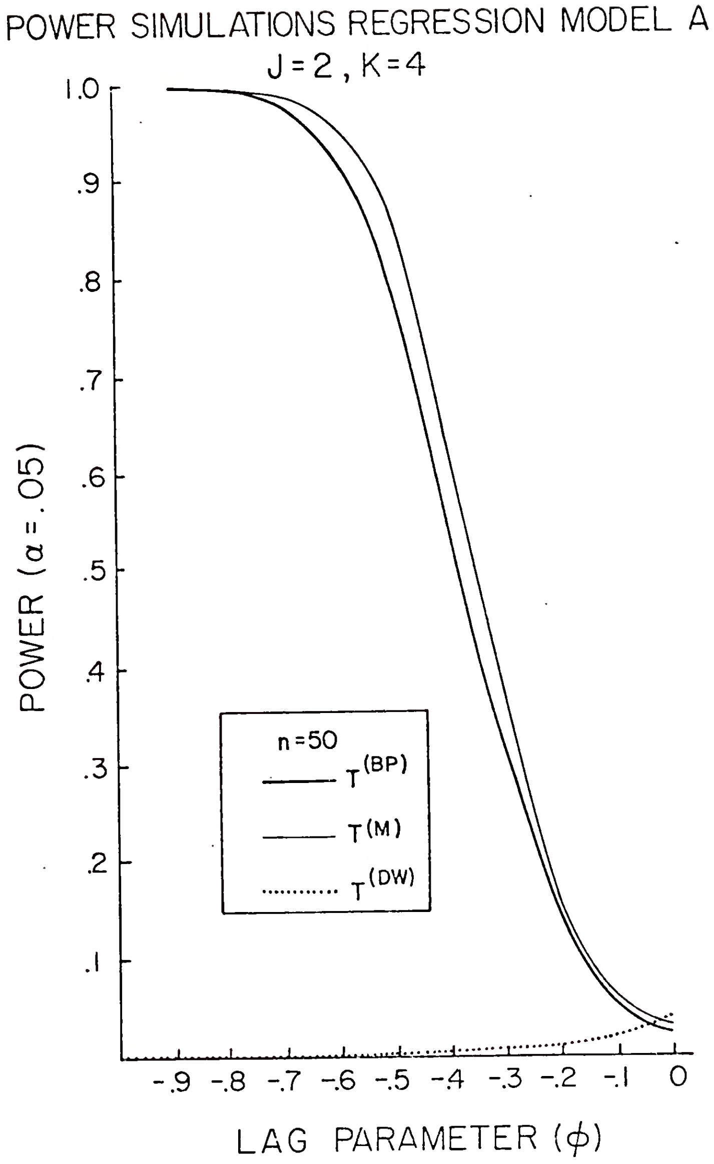

POWER SIMULATIONS REGRESSION MODEL A
J=2, K=4
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
n=50
$T^{(BP)}$
$T^{(M)}$
$T^{(DW)}$
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (ϕ)

FIGURE 4.6a

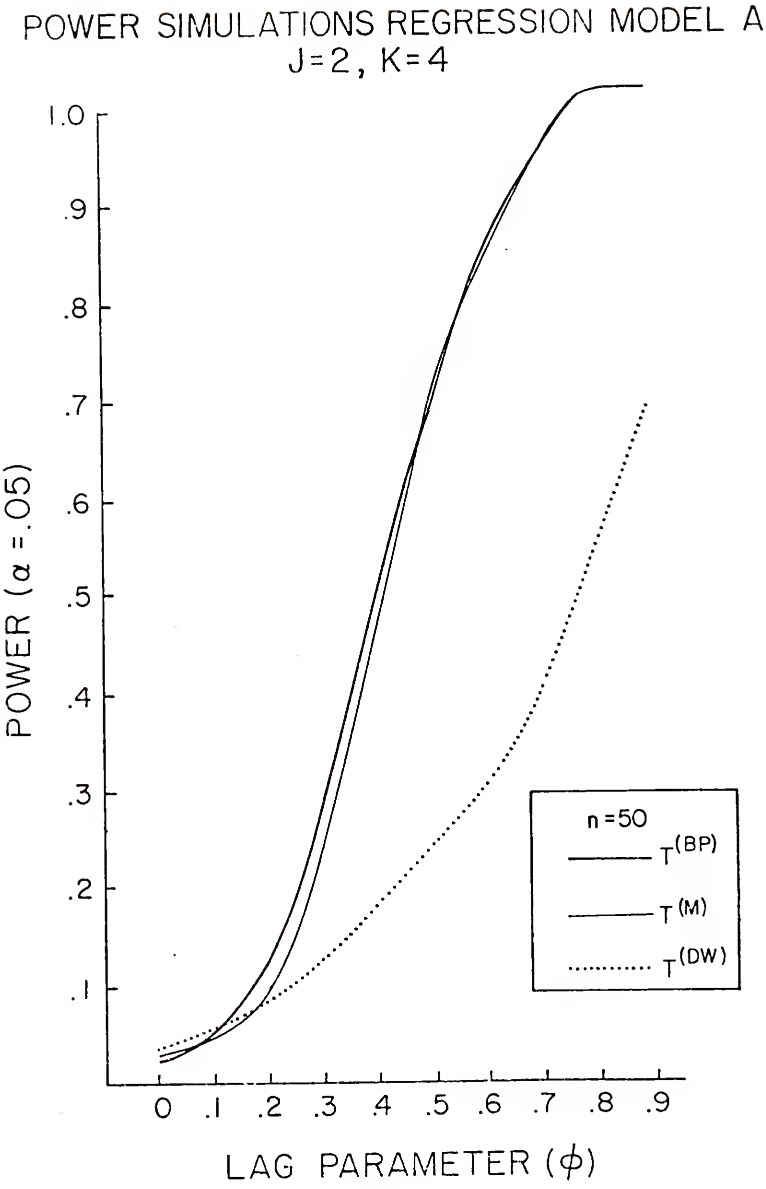

POWER SIMULATIONS REGRESSION MODEL A
J=2, K=4
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
POWER (α = .05)
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (ϕ)
n=50
$T^{(BP)}$
$T^{(M)}$
$T^{(DW)}$

FIGURE 4.6b

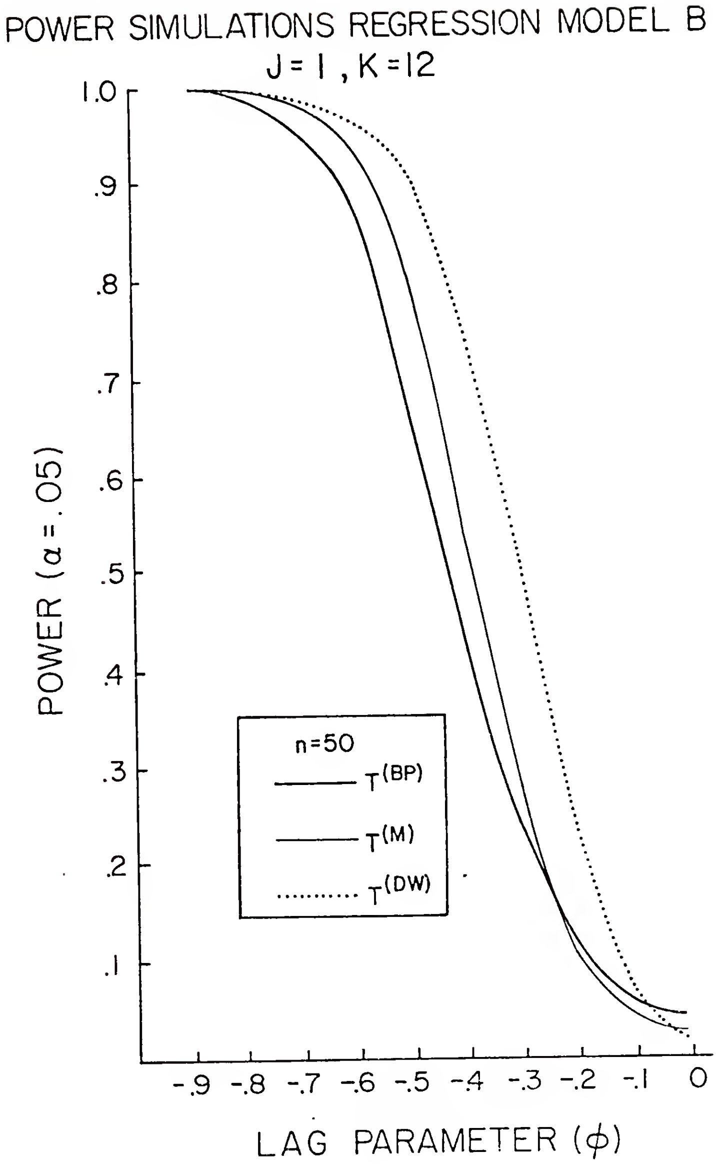

POWER SIMULATIONS REGRESSION MODEL B
J = 1 , K = 12
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
n=50
$T^{(BP)}$
$T^{(M)}$
$T^{(DW)}$
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (ϕ)

FIGURE 4.7a

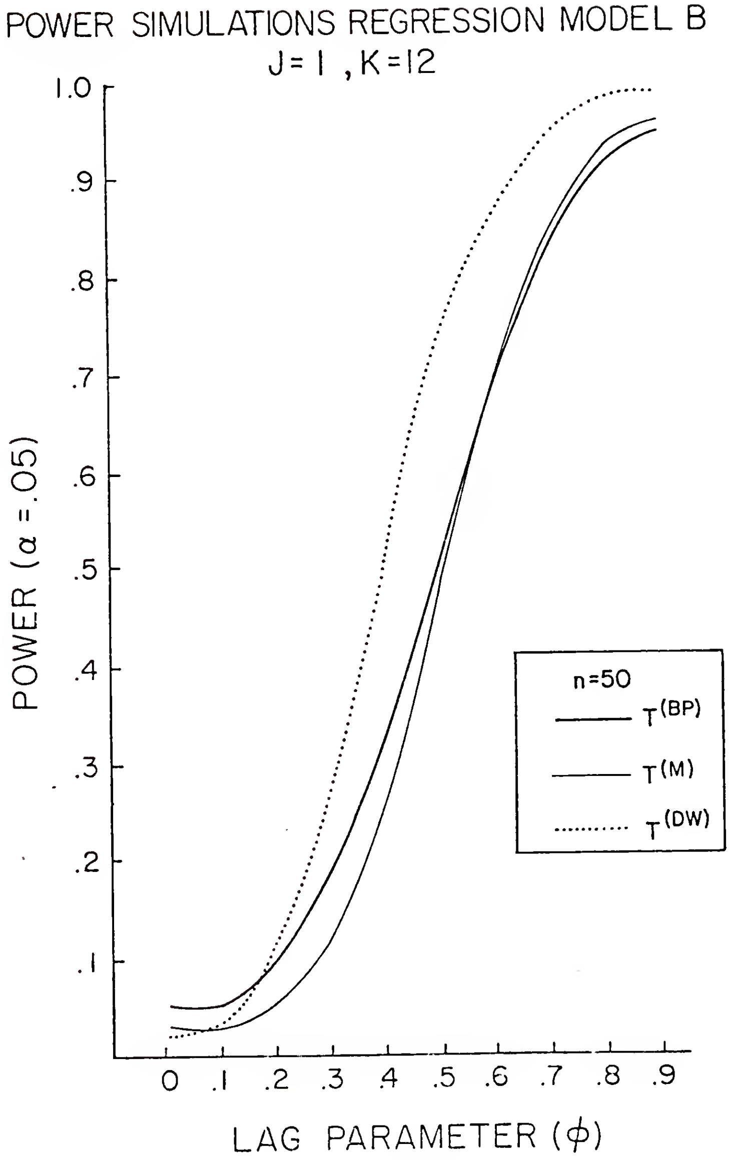

POWER SIMULATIONS REGRESSION MODEL B
J = 1 , K = 12
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
POWER (α = .05)
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (ϕ)
n=50
$T^{(BP)}$
$T^{(M)}$
$T^{(DW)}$

FIGURE 4.7b

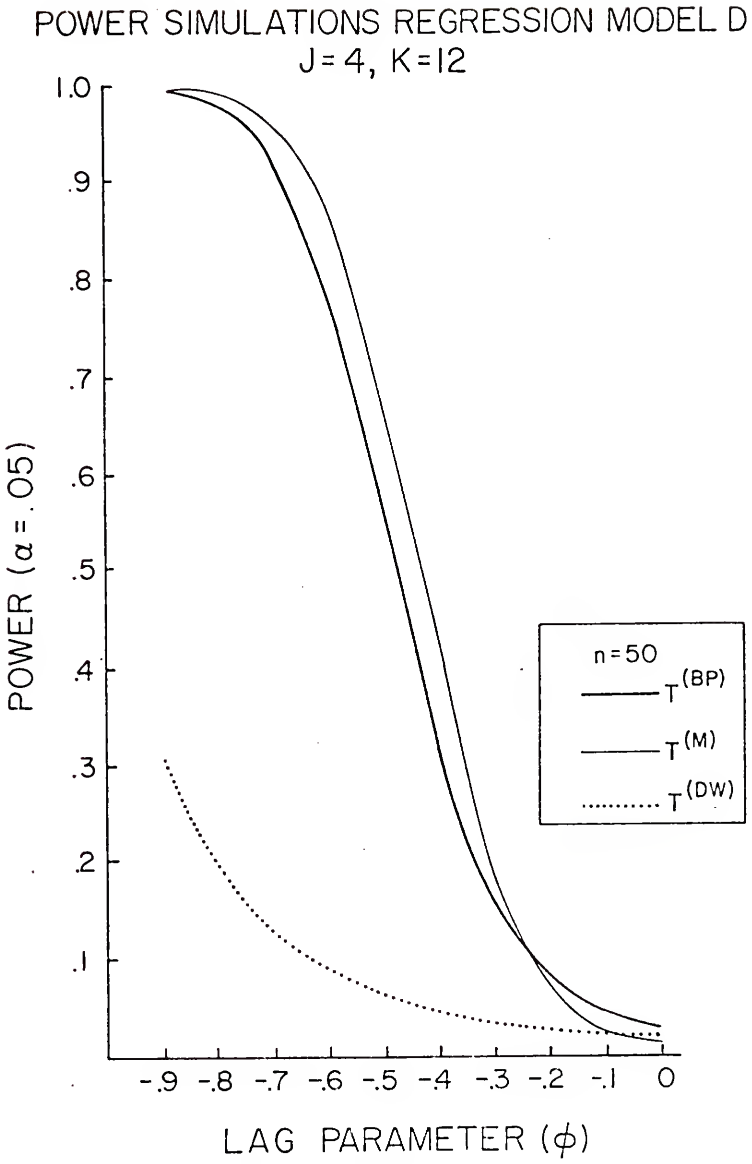

POWER SIMULATIONS REGRESSION MODEL D
J = 4, K = 12
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (ϕ)
n = 50
$T^{(BP)}$
$T^{(M)}$
$T^{(DW)}$

FIGURE 4.8a

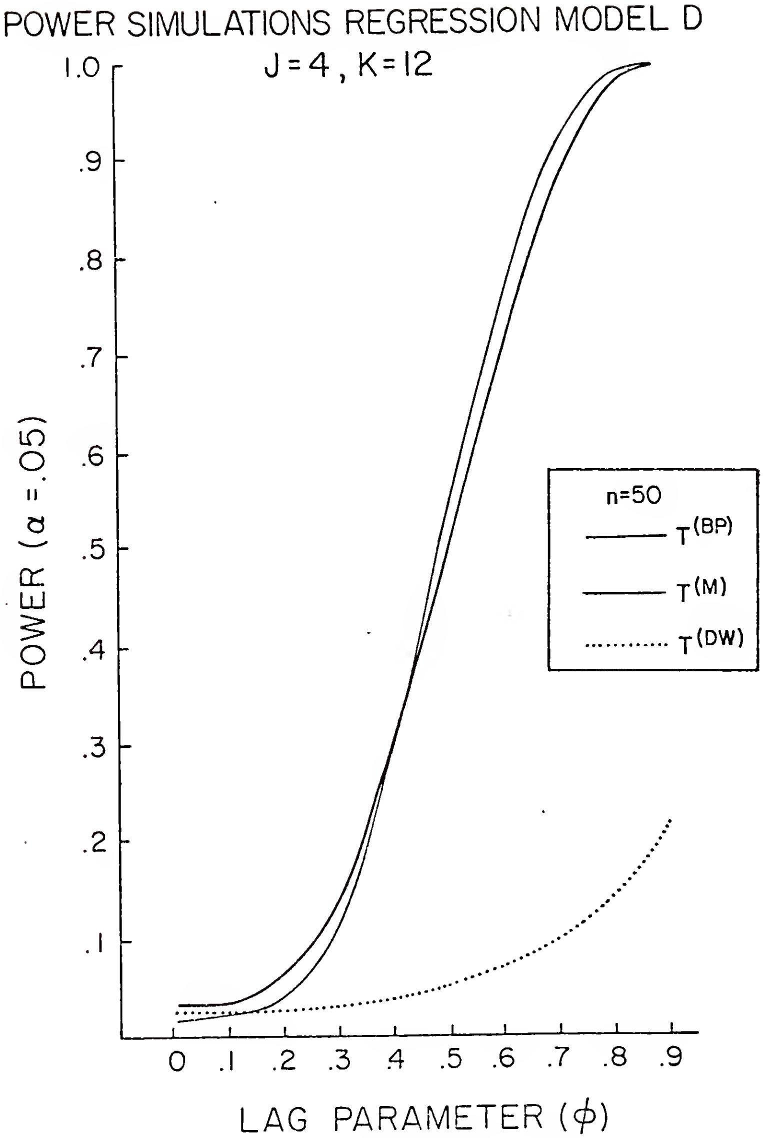

POWER SIMULATIONS REGRESSION MODEL D
J=4, K=12
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (φ)
n=50
T(BP)
T(M)
T(DW)

FIGURE 4.8b

The Durbin-Watson d test offers a very intersting set of results. While the power curves for the Box-Pierce and max-χ^2 tests depicted in Figures 4.6 - 4.8 remain virtually unchanged, the power curves for the d test are very diversified. Consider first the case J=1, K=12 (Figure 4.7). In this instance, the power of the d test clearly dominates the powers of both the other tests for all values of ϕ. This is to be expected, however, since the d statistic essentially reduces to r_1, the sample first order autocorrelation function (Section 3.1). We have seen that r_1 is sufficient for ρ_1, and hence ϕ (Lemma 15). Figure 4.8 displays the tremendous loss in the d test's power when J increases to 4. For almost all values of ϕ in the case J=4, K=12, the power of the d test is .1 or lower, while the maximum power attained by the test is .3 (see Table A-28 of the Appendix). However, it is when J=2 (Figure 4.6) that the power trend of the d test makes its most dramatic change. For every other case, the power of the d test increases as the lag parameter ϕ increases in absolute value. In the case J=2, K=4, though, the d test has essentially no power (power close to zero) for large negative values of ϕ and attains an atypical high power of .703 when ϕ = .9 (Table A-13). This abrupt change in the power pattern can be partially explained by investigating the asymptotic distribution of $\sqrt{n}\, r_1$ when J=2. Recall from Theorem 1 that the variance of the asymptotic distribution of $\sqrt{n}\, r_1$ is $v_{11} = \frac{(1+\phi)^2}{1-\phi^2}$. As ϕ increases, v_{11} also increases. Hence for large negative values of the lag parameter, the asymptotic variance of $\sqrt{n}\, r_1$ will be small, and since $\rho_1 = 0$ for this case, $\sqrt{n}\, r_1$ or d will vary over a small range centered at 0. Thus, the computed value of d will rarely fall into the critical region when ϕ is large and negative. However, for large positive ϕ, the reverse is true. The d statistic will vary over a large range centered at 0, leading to larger computed values and a higher rejection rate. A quick glance at the last column of Tables A-13 - A-28 in the Appendix reveals that the rate at which the d statistic falls into the "inconclusive" region is relatively small for all cases. This, again, is to be expected, since n-g is large (either 48 or 49, depending on the regression model).

Overall, the different regression models have very little effect on the general power trends of the tests. However, a comparison of the results of regression reveals the following: for negative values of the lag parameter, the powers of the Box-Pierce and max-χ^2 tests are generally higher when the test statistics are calculated using regression residuals; for positive values of the lag parameter the reverse is true, that is, the powers of the tests are generally higher when the test statistics are calculated using the "observable" residuals. This result is due in part to the bias introduced by the application of the least squares regression and is best explained in the following sequential manner:

(i) For positive values of $\emptyset$, the actual variability of the least squares estimator $\hat{\underset{\sim}{\beta}}$ is larger than indicated by the least squares variance formulae (see Section 1.1).

(ii) Since the vector of least squares residuals $\hat{\underset{\sim}{Z}}$ is computed from $\hat{\underset{\sim}{\beta}}$, the residual variances are larger than would be expected in the white noise case.

(iii) As a consequence of (ii), the sample autocorrelations computed from the least squares residuals also tend to have larger variances.

(iv) The Box-Pierce and max-χ^2 statistics have non-null distributions centered away from 0 when $\emptyset$ is nonzero. Thus, the tests will fail to reject the hypothesis of white noise ($\phi = 0$) when in fact the alternative is true ($\phi \neq 0$) only if random variation causes the test statistics to fall in their respective lower tail areas.

(v) The larger variability of the sample autocorrelations (ii), of course, increase the variances of both the Box-Pierce and max-χ^2 test statistics. The chance of random variation giving rise to "small" values of the test statistics is in turn greater, which leads to a less powerful test.

For negative values of ϕ, (i) through (v) hold in reverse, and the tests become more powerful in the regression case.

The main results of the simulation study for the regression case are summarized as follows:

(i) The power trends detected for regression models A, B, C and D are essentially the same.

(ii) The power of the max-χ^2 test dominates the power of the Box-Pierce test for all negative and moderate to large positive values of the lag parameter ϕ.

(iii) The power of the Box-Pierce test is slightly larger than the power of the max-χ^2 test when ϕ is positive and small, although both tests have low power.

(iv) The Durbin-Watson d test outperforms both the Box-Pierce and max-χ^2 tests under a first order lag 1 alternative.

(v) Both the Box-Pierce and max-χ^2 tests outperform the d test for those first order alternatives with lag larger than 1.

Since the first order AR scheme with lag 1 represents a very common and hence, a very important residual error model in practice, and since the

Durbin-Watson d test attains a very high power in this case, we recommend that the practical researcher always apply the Durbin-Watson procedure as a test for serial correlation of the errors of a regression model. However, the d test should be supplemented by another test which has high power against more general alternative error models. From our simulation study, we have seen that both the Box-Pierce and max-χ^2 tests make excellent candidates. But the max-χ^2 test appears to be the better supplemental test if a low order model is expected, for our results consistently show the dominance of the max-χ^2 power over the Box-Pierce power in this case.

4.2 Power Approximations for Large Samples

Our final goal of the simulation study is to compare the simulated powers of the Box-Pierce and max-χ^2 tests to the large sample power approximations derived in Section 3.3. Recall that our method of approximating the max-χ^2 power was to use a Taylor-series expansion of the asymptotic joint density function of the first K sample autocorrelations. The expansion is given in equation (3.3.27) of Lemma 12. We considered three different methods for approximating the power of the Box-Pierce test: (a) Imhoff numerical integration, (b) Imhoff extension to the Pearson three-moment approximation, and (c) Rao asymptotic normal square root transformation. These approximations to the Box-Pierce power, $P\{T^{(BP)} > C_\alpha^{(BP)}\}$, are expressed as follows (for details, see Section 3.3.3.2.1):

a. $\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\sin\theta(u)}{u\rho(u)}\,du$, where the functions $\theta(u)$ and $\rho(u)$ are given by equation (3.3.18) of Theorem 7.

b. $P\{\chi_h^2(0) > C^*\}$, where h and C* are defined in equation (3.3.19) of Theorem 8.

c. $$1-\Phi\left\{\frac{\sqrt{c_\alpha^{(BP)}} - \sqrt{n}(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{p}'V\underset{\sim}{\rho})^{\frac{1}{2}}}\right\} + \Phi\left\{\frac{-\sqrt{c_\alpha^{(BP)}} - \sqrt{n}(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{\frac{1}{2}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}}\right\} ,$$

where $\underset{\sim}{\rho}$ and V are the asymptotic mean vector and asymptotic covariance matrix, respectively, of the first K sample autocorrelations.

For this part of the study, we consider three versions of error model 1: (J=2, K=4), (J=3, K=4) and (J=4, K=4). In order to apply the approximations, we must derive the asymptotic mean vector $\sqrt{n}\ \underset{\sim}{\rho}$ and covariance matrix V of the first four sample autocorrelations in each case. Recall from Theorem 2 of Section 3.3.3 that the eigenvalues of V, their respective orders of multiplicities and certain linear combinations of the elements of $\underset{\sim}{\rho}$ all contribute to the power expressions for the Box-Pierce approximations and max-χ^2 approximations. In the following subsections, we give $\sqrt{n}\ \underset{\sim}{\rho}$ and V for each of the three cases, and also the Taylor-series approximation to the max-χ^2 power. We then use these results to compute the power approximations.

4.2.1 Case: J=4, K=4

We initially considered the case (J=2, K=4) in Section 3.3.3.2. Recall that the asymptotic mean vector and covariance matrix of $\sqrt{n}\ \underset{\sim}{r} = \sqrt{n}\ (r_1, r_2, r_3, r_4)'$ are as follows:

$$\sqrt{n}\ \underset{\sim}{\rho} = \sqrt{n}\ (0, \emptyset, 0, \emptyset^2)' \quad ,$$

$$V = \begin{bmatrix} \frac{(1+\emptyset)^2}{1-\emptyset^2} & 0 & \frac{2\emptyset+3\emptyset^2-\emptyset^4}{1-\emptyset^2} & 0 \\ 0 & 1-\emptyset^2 & 0 & 2\emptyset(1-\emptyset\ ) \\ \frac{2\emptyset+3\emptyset^2-\emptyset^4}{1-\emptyset^2} & 0 & 1+\emptyset^2+4\emptyset^3-2\emptyset^5 & 0 \\ 0 & 2\emptyset(1-\emptyset^2) & 0 & 1+\emptyset^2-3\emptyset^4 \end{bmatrix}$$

The Taylor-series approximation to the max-χ^2 power is given by equation (3.3.29) of Theorem 9.

4.2.2 Case: J=3, K=4

Substitute J=3 in equation (3.3.4) of Lemma 4 and equations (3.3.9) and (3.3.10) of Theorem 4 to obtain the asymptotic mean vector and covariance matrix of $\sqrt{n}\,\underset{\sim}{r}$:

$$\sqrt{n}\,\underset{\sim}{\rho} = \sqrt{n}\ (0,\ 0,\ \emptyset,\ 0)' ,$$

$$V = \begin{bmatrix} \frac{1+\emptyset^2}{1-\emptyset^2} & \frac{2\emptyset}{1-\emptyset^2} & 0 & \frac{2\emptyset}{1-\emptyset^2} \\ \frac{2\emptyset}{1-\emptyset^2} & \frac{1+\emptyset^2}{1-\emptyset^2} & 0 & \frac{3\emptyset^2-\emptyset^4}{1-\emptyset^2} \\ 0 & 0 & 1-\emptyset^2 & 0 \\ \frac{2\emptyset}{1-\emptyset^2} & \frac{3\emptyset^2-\emptyset^4}{1-\emptyset^2} & 0 & \frac{1+\emptyset^2}{1-\emptyset^2} \end{bmatrix}$$

The Taylor-series expansion approximation to the max-χ^2 power is, for this case, given by the following theorem.

Theorem 11: A Taylor-series approximation to the asymptotic power of the max-χ^2 test for the case (J=3, K=4) is obtained by computing:

$$1-[\Phi(b_3)-\Phi(a_3)]\cdot\{[2\Phi(b_1)-1]^3+\frac{1}{2}[2\Phi(b_1)-1][4b_1^2\phi^2(b_1)][2C_1^2+C_2^2] \tag{4.2.1}$$

$$+\frac{1}{6}[-8b_1^3\phi^3(b_1)][C_1^2C_2]+\frac{1}{24}([2\Phi(b_1)-1][4(b_1^3-3b_1)^2\phi^2(b_1)][2C_1^4+C_2^4]$$

$$+[-8(b_1^3-3b_1)b_1^2\phi^3(b_1)][C_1^4+2C_1^2C_2^2])+\frac{1}{120}[-8(b_1^3-3b_1)b_1^2\phi^3(b_1)][2C_1^4C_2+C_1^2C_2^3]$$

$$+\frac{1}{720}([2\Phi(b_1)-1][4(b_1^5-10b_1^3+15b_1)^2\phi^2(b_1)][2C_1^6+C_2^6]$$

$$+[-8(b_1^5-10b_1^3+15b_1)(b_1^3-3b_1)b_1\phi^3(b_1)][2C_1^6+2C_1^4C_2^2+2C_1^2C_2^4]$$

$$+[-8(b_1^3-3b_1)^3\phi^3(b_1][C_1^4C_2^2])\}\quad ,$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, $\Phi(x) = \int_{-\infty}^{x} \phi(z)dz$, $C_1 = \frac{2\emptyset}{1+\emptyset^2}$, $C_2 = \frac{3\emptyset 2-\emptyset^4}{1+\emptyset^2}$,

$$b_1 = \frac{\sqrt{c^{(BP)}(1-\emptyset^2)}}{\sqrt{1+\emptyset^2}} , \quad b_3 = \frac{\sqrt{c_\alpha^{(BP)}} - \sqrt{n}\,\emptyset}{\sqrt{1-\emptyset^2}} , \quad a_3 = \frac{-\sqrt{c_\alpha^{(BP)}} - \sqrt{n}\,\emptyset}{\sqrt{1-\emptyset^2}} .$$

Proof: Follows the proof of Theorem 9. □

4.2.3 Case: J=4, K=4

Apply Lemma 4 and Theorem 4 to obtain

$$\sqrt{n}\,\underset{\sim}{\rho} = \sqrt{n}(0,\ 0,\ 0,\ \emptyset)' ,$$

$$V = \begin{bmatrix} \frac{1+\emptyset^2}{1-\emptyset^2} & 0 & \frac{2\emptyset}{1-\emptyset^2} & 0 \\ 0 & \frac{(1+\emptyset)^2}{1-\emptyset} & 0 & 0 \\ \frac{2\emptyset}{1-\emptyset^2} & 0 & \frac{1+\emptyset^2}{1-\emptyset^2} & 0 \\ 0 & 0 & 0 & 1-\emptyset^2 \end{bmatrix}$$

the mean vector and covariance matrix, respectively, of the asymptotic distribution of $\sqrt{n}\,\underset{\sim}{r} = \sqrt{n}(r_1, r_2, r_3, r_4)'$ when J=4. From Lemma 12, we derive the Taylor-series approximation to the max-χ^2 power, and give the result as a theorem.

Theorem 12: A Taylor-series approximation to the asymptotic power of the max-χ^2 test for the case (J=4, K=4) is obtained by computing:

$$1-[\Phi(b_4)-\Phi(a_4)][2\Phi(b_2)-1]\{[2\Phi(b_1)-1]^2+ \frac{1}{2}[C^2][4b_1^2\phi^2(b_1)] \qquad (4.2.2)$$

$$+ \frac{1}{24}[4(b_1^3-3b_1)^2\phi^2(b_1)C^4] + \frac{1}{720}[4(b_1^5-10b_1^3+15b_1)^2\phi^2(b_1)][C^6]$$

$$+ \frac{1}{8!}[4(b_1^7-21b_1^5+105b_1^3-105b_1)^2\phi^2(b_1)][C^8]\} ,$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, $\Phi(x) = \int_{-\infty}^{x} \phi(Z)dz$, $C = \frac{2\emptyset}{1+\emptyset^2}$,

$$b_1 = \frac{\sqrt{c_\alpha^{(BP)}(1-\phi^2)}}{\sqrt{1+\phi^2}}, \quad b_2 = \frac{\sqrt{c_\alpha^{(BP)}(1-\phi^2)}}{1+\phi}, \quad b_4 = \frac{\sqrt{c_\alpha^{(BP)}} - \sqrt{n}\phi}{\sqrt{1-\phi^2}}$$

$$a_4 = \frac{-\sqrt{c^{(BP)}} - \sqrt{n}\phi}{\sqrt{1-\phi^2}} .$$

Proof: Follows the proof of Theorem 9. □

4.2.4 Approximation Results

We computed the power approximations of both the Box-Pierce and max-χ^2 tests for various values of the lag parameter ϕ and the sample size n in each of the three versions of error model 1 discussed above. In order to determine the accuracy of the approximations, we again simulated the true power of the tests (using the "observable" residuals as in Section 4.1.1). The parameters of the study are given in Table 4.7. Once again, we sketch those power curves which best illustrate the power approximation results, and refer the reader to Appendix for tables of numerical values.

The results for all three cases are printed in Tables A-29 - A-34 of the Appendix. Since the results were so similar, we only sketch power curves for the J=2, K=4 case. Consider first the n=50 results for the Box-Pierce approximations (Figure 4.9). Generally, the numerical integration (BP1) and Pearson extension (BP2) approximations tend to overestimate the simulated power of the Box-Pierce test, while the Rao square root transformation (BP3) tends to underestimate the power. At large values of the lag parameter, BP3 gives a slightly better approximation than BP1. For moderate values of ϕ, i.e., $.3 \leq |\phi| \leq .7$, the simulated Box-Pierce power is bounded above by BP1 and below by BP3. For ϕ in this range, BP2 performs poorly. The BP2 approximation seems

TABLE 4.9

POWER APPROXIMATIONS FOR THE BOX-PIERCE AND MAX-χ^2 TESTS
ERROR MODEL 1: $Z_t = \emptyset Z_{t-J} + \varepsilon_t$

Parameters:	(J,K)	Ø	n	α
	(2,4)	-.9	50	.05
	(3,4)	-.8	200*	
	(4,4)	-.7		
		-.6		
		-.5		
		-.4		
		-.3		
		-.2		
		-.1		
		0		
		.1		
		.2		
		.3		
		.4		
		.5		
		.6		
		.7		
		.8		
		.9		

*NOTE: For n=200, powers are approximated only for odd values of Ø

TABLE 4.10

KEY TO SYMBOLS USED IN POWER APPROXIMATION FIGURES

Box-Pierce: BP(1) = Imhoff numerical integration

BP(2) = Imhoff's extension to the Pearson 3-moment approximation

BP(3) = Rao square root transformation

BP(4) = Simulated power (no regression)

Max-χ^2 : M(1) = Taylor-series approximation

M(2) = Simulated power (no regression)

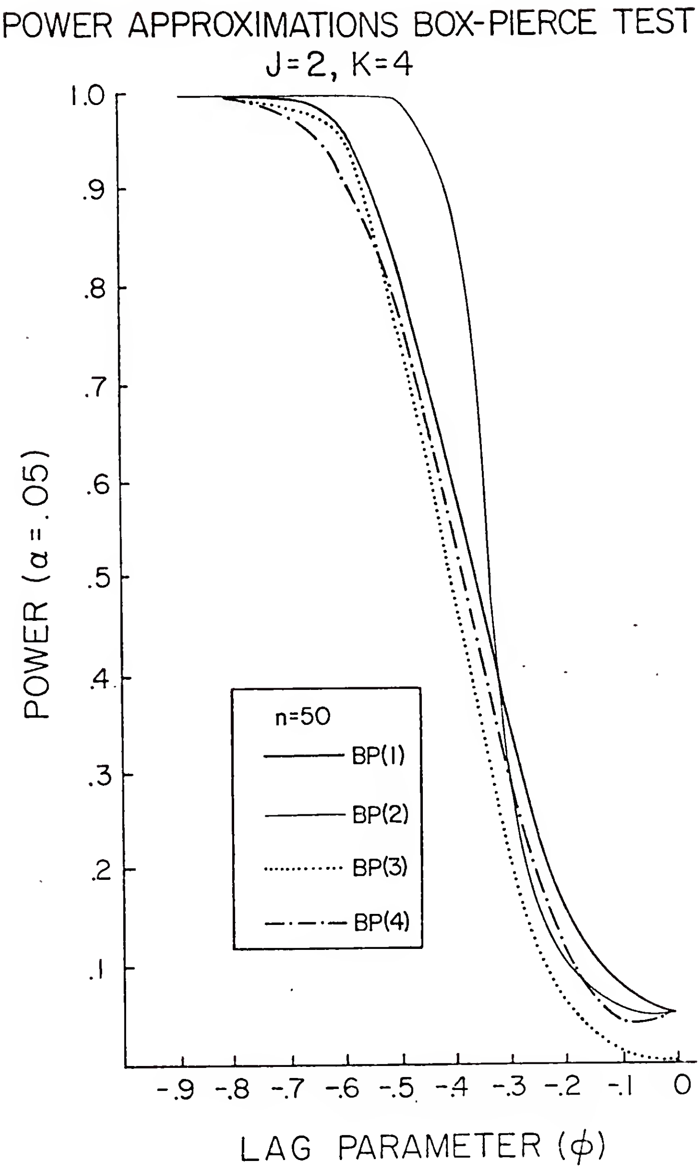

POWER APPROXIMATIONS BOX-PIERCE TEST
J=2, K=4
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
n=50
BP(1)
BP(2)
BP(3)
BP(4)
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (φ)

FIGURE 4.9a

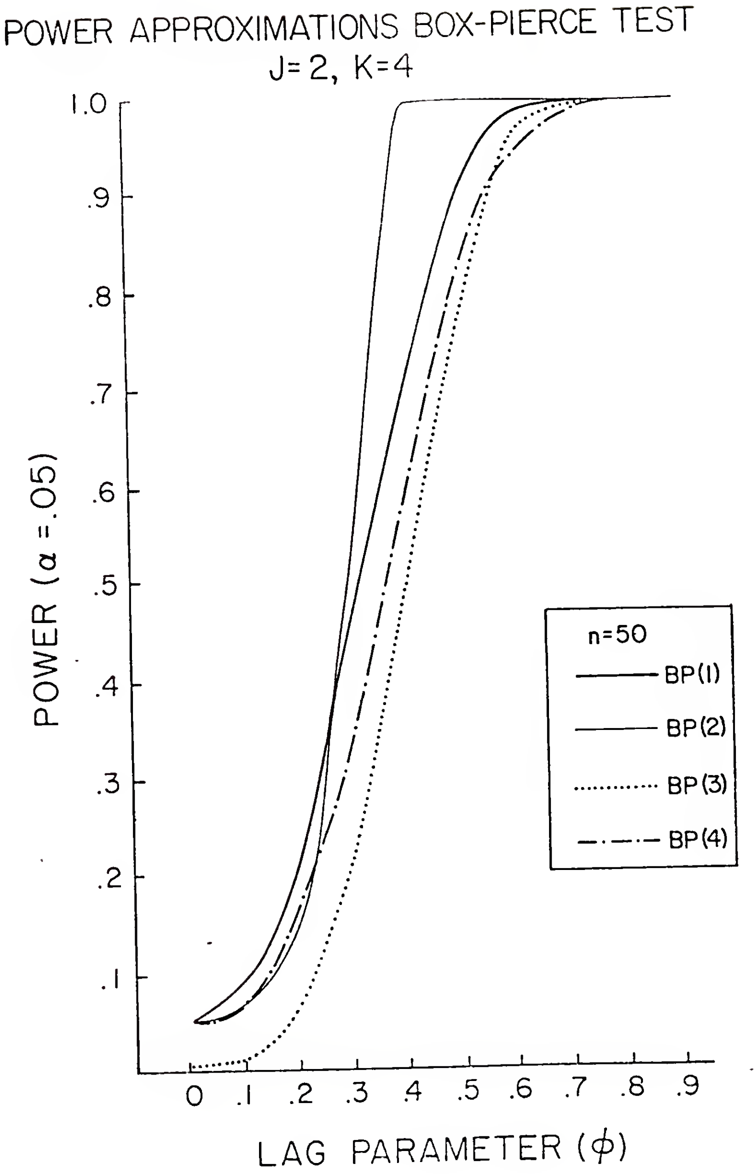

POWER APPROXIMATIONS BOX-PIERCE TEST
J=2, K=4
POWER (α =.05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (φ)
n=50
BP(1)
BP(2)
BP(3)
BP(4)

FIGURE 4.9b

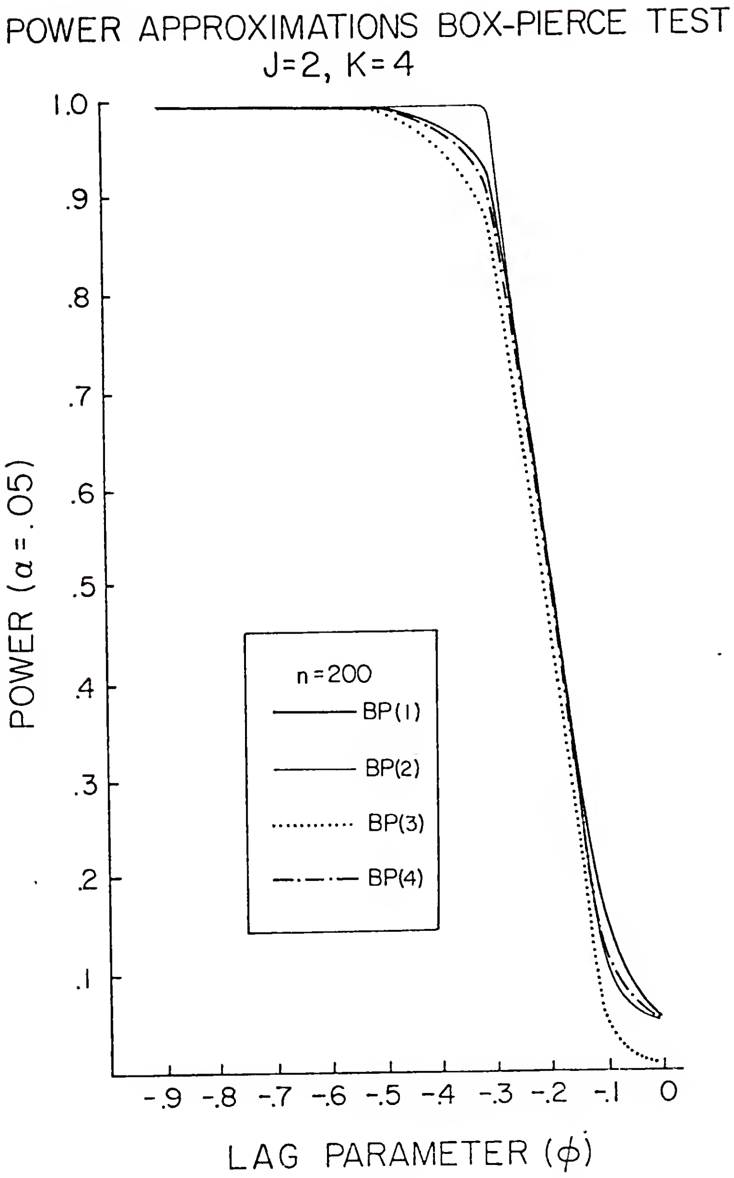

POWER APPROXIMATIONS BOX-PIERCE TEST
J=2, K=4
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
POWER (α = .05)
n=200
BP(1)
BP(2)
BP(3)
BP(4)
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (ϕ)

FIGURE 4.10a

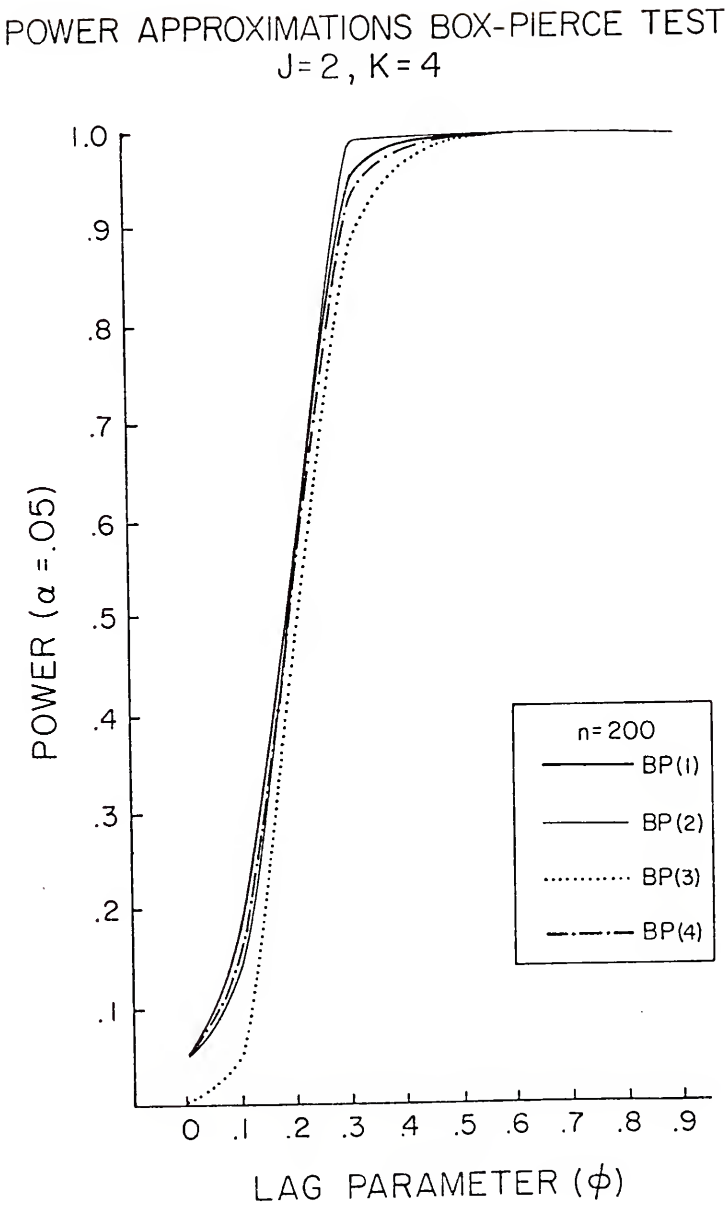

POWER APPROXIMATIONS BOX-PIERCE TEST
J=2, K=4
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
POWER (α = .05)
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (φ)
n=200
BP(1)
BP(2)
BP(3)
BP(4)

FIGURE 4.10b

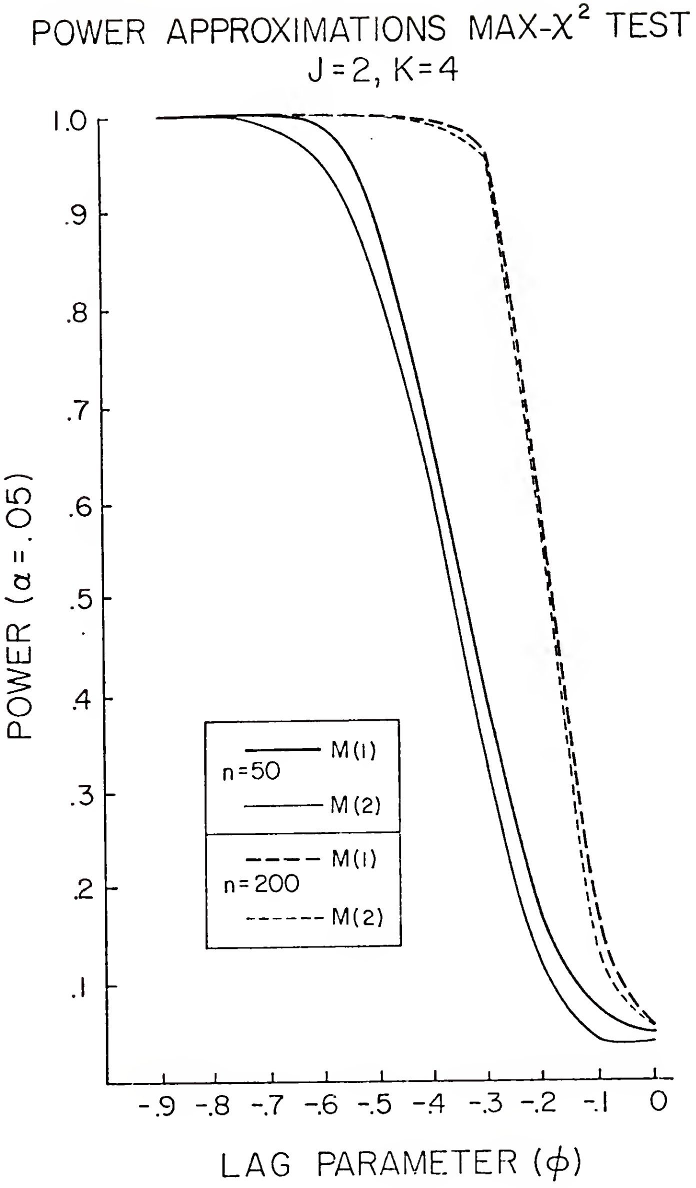

POWER APPROXIMATIONS MAX-X² TEST
J=2, K=4
POWER (α=.05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
M(1)
n=50
M(2)
M(1)
n=200
M(2)
-.9
-.8
-.7
-.6
-.5
-.4
-.3
-.2
-.1
0
LAG PARAMETER (ϕ)

FIGURE 4.11a

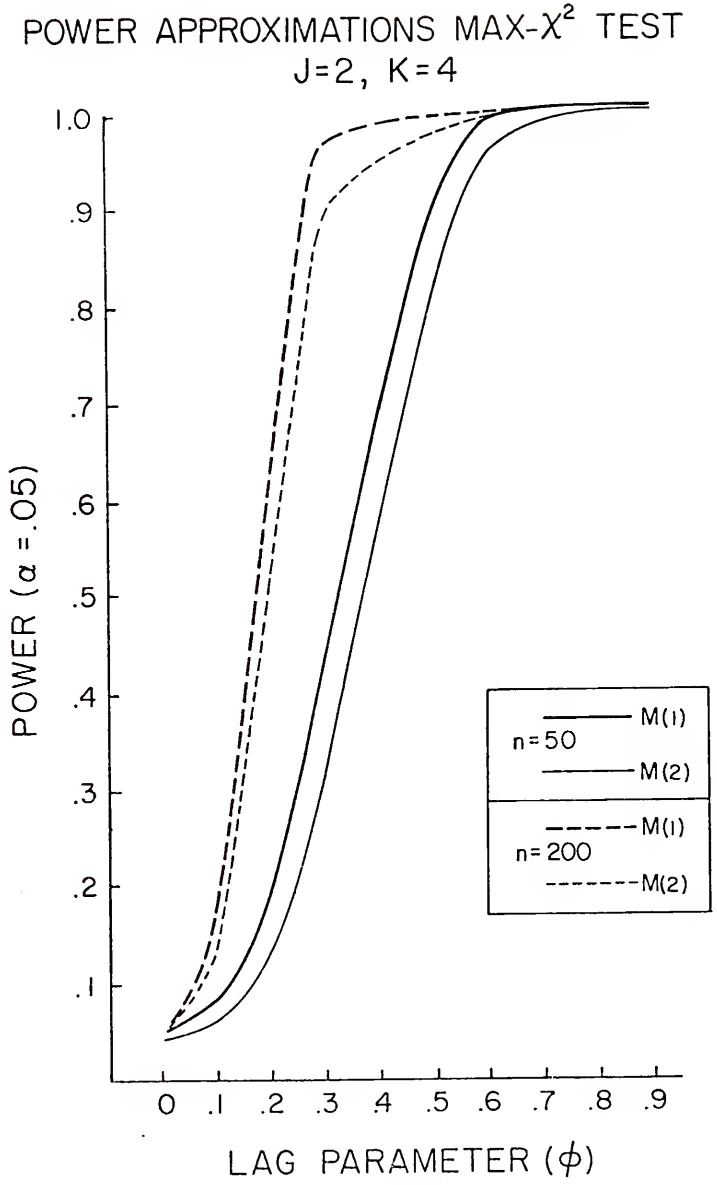

POWER APPROXIMATIONS MAX-χ^2 TEST
J=2, K=4
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
LAG PARAMETER (ϕ)
n=50
M(1)
M(2)
n=200
M(1)
M(2)

FIGURE 4.11b

to be the most accurate when the lag parameter is small. Overall, it appears that the Rao square root transformation (BP3) gives the more accurate results when n=50. Notice, however, that when $-.1 \leq \phi \leq .1$, BP3 gives an approximation which is very nearly 0. The attained significance level of Rao's procedure, then, is much smaller than .05. From our simulation study in Section 4.1.1, we discovered that the actual level of the test (simulated) for n=50 is less than the asymptotic level (.05). The Rao square root transformation appears to work well at n=50 because it also underestimates the asymptotic level, and hence, expected power of the Box-Pierce test using the asymptotic distribution. Figure 4.10 shows that when n is increased to 200, BP3 does not perform well. For n this large, the asymptotic power of the Box-Pierce test is nearly attained, hence the simulated power better estimates the asymptotic power, and the bias in the BP3 approximation is more readily observed. For n=200, Imhoff's numerical integration procedure (BP1) appears to give an excellent approximation to the power. This result should not be surprising, for one must recall that Imhoff's procedure is only an "approximation" in the sense that the numerical integration technique estimates the true integral given in the power expression, (Section 3.3.3.2). Imhoff derives the result by inverting the characteristic function of the Box-Pierce test statistic under the asymptotic distribution. The procedure would be exact asymptotically if the integral could be evaluated analytically. Consequently, when n is large enough for the simulated power to closely approximate the asymptotic power of the test, as is the case when n=200, a comparison of BP1 to BP4 gives results which favor Imhoff's technique. Thus, it is very likely that at n=50, the numerical integration procedure gives the best approximation to the asymptotic power of the Box-Pierce test.

From these results, we write the Box-Pierce approximations in the following increasing order of accuracy: (1) Imhoff's extension to the Pearson 3-moment approximation; (2) Rao asymptotic normal square root transformation, and (3) Imhoff numerical integration. For those researchers without a computer routine, we note that the approximations are also listed in increasing order of difficulty of computation.

Refer now to Figure 4.11 in order to gauge the accuracy of the Taylor-series approximation (M1) to the max-χ^2 power. The results indicate that the approximation overestimates the simulated power of the test, although at n=200, the bias is small. Again, we can attribute the apparently less accurate results for n=50 to the fact that the simulated power, using the critical region given by the asymptotic approximation to the tail area of the max-χ^2 distribution, underestimates the asymptotic power of the test. The magnitude of ϕ (in absolute value) does not appear to affect the overall accuracy of the approximation. We warn that the researcher without access to computer facilities will encounter many tedious and difficult calculations when using this approximation (Section 4.2.5). However, the Taylor-series expansions are easily programmed.

We present a summary of the power approximation results:

(i) The accuracy of the power approximations is not affected by the change in the lag (J) of the error model.

(ii) Imhoff's extension to Pearson's 3-moment approximation performs poorly as an estimate of the true Box-Pierce power when the lag parameter is moderate, i.e., $3 \leq |\phi| \leq .7$.

(iii) Rao's asymptotic normal square root transformation tends to underestimate the asymptotic Box-Pierce power, especially when ϕ is small.

(iv) Imhoff's numerical integration technique is the most accurate among the Box-Pierce power approximations for large n.

(v) The Taylor-series approximation tends to overestimate the asymptotic max-χ^2 power, although this bias decreases as n is increased.

4.2.5 A Note on the Taylor-Series Approximation

We have stated that the researcher who attempts to calculate the Taylor-series approximation to the max-χ^2 power without the aid of a computer will face a difficult task, for the computations are quite tedious. As part of an investigation into simplifying the approximation, we computed each term of the series expansion separately, in order to gauge its contribution to the overall estimated power. Recall from Section 3.3.3.2.2. the expression for the asymptotic power of the max-χ^2 test for K=4:

$$P\{T^{(m)} > C_\alpha^{(m)}\} = 1 - \int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3}\int_{a_y}^{b_y} f(\underset{\sim}{y}\colon \underset{\sim}{0}, R)\, dy_1 dy_2 dy_3 dy_4 \qquad (4.2.3)$$

where $\{a_i\}$ and $\{b_i\}$ are constants which depend upon $C_\alpha^{(m)}$, $\underset{\sim}{\rho}$ and V, f is the asymptotic joint density of the first four sample autocorrelations, and R is the correlation matrix corresponding to the covariance matrix V. We approximated the integral in equation (4.2.3) by a Taylor-series expansion. Notationally, let us write the max-χ^2 power (M_1) as follows:

$$M_1 = 1 - \{T_o + T_1 + T_2 + . . .\} = 1 - M_o$$

where T_o, T_1, T_2, . . . are the power contributions for each term in the series expansion and

$$M_o = \sum_{i=0}^{\infty} T_i \quad .$$

Computations for all three cases discussed in this section, i.e., (J=2, K=4), (J=3, K=4) and (J=4, K=4) for samples of size 50 and 200 are listed in Tables A-35 - A-40 of the Appendix. Here, we give only the percent contribution of T_o to the sum of all terms M_o in the Table 4.11.

Except for the case J=2, K=4, n=50, $|\phi| = .7$, the ratio of the initial term T_o in the Taylor-series approximation to the sum of the terms M_o is at worst .7 and in most cases, very nearly 1. Thus, the user of the Taylor-series expansion technique can simplify calculations considerably by only considering the first term of the expansion, and still obtain a rather accurate estimate of the max-χ^2 power. To obtain T_o, we let t=0 in equation (3.2.23) which gives the first term of the series expansion of the asymptotic joint density of the first K sample autocorrelations. This amounts to substituting the identity matrix I for R in (4.2.3). Hence, our simplified approximation of the power is given as:

$$P\{T^{(m)} > C_\alpha^{(m)}\} \approx 1 - \int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3}\int_{a_4}^{b_4} f(\underset{\sim}{y}: 0, I)\, dy_1 dy_2 dy_3 dy_4 \qquad (4.2.4)$$

Now the right-hand side of (4.2.4) is easily integrated. Using Theorem 10, we obtain

$$P\{T^{(m)} > C_\alpha^{(m)}\} \approx 1 - \prod_{i=1}^{4} [\Phi(b_i) - \Phi(a_i)] \qquad (4.2.5)$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$ and $\{b_i\}$, $\{a_i\}$ are given by equation (3.3.29). The researcher only needs to have access to tables of the cumulative normal distribution function, which are readily available. The point to be made here is that even though the covariance matrix V of the first K sample autocorrelations, $r_1, r_2, \ldots, r_K$, is non-diagonal,

TABLE 4.11

TAYLOR-SERIES APPROXIMATION CONTRIBUTION OF T_o TO THE SUM M_o*

	(J=2, K=4)		(J=3, K=4)		(J=4, K=4)	
∅	n=50	n=200	n=50	n=200	n=50	n=200
-.9	--	--	--	--	--	--
-.7	.278	--	.747	--	.886	--
-.5	.741	--	.949	--	.974	--
-.3	.971	.878	.993	.993	.997	.997
-.1	.999	.999	.999	.999	.999	.999
.1	.999	.998	.999	.999	.999	.999
.3	.962	.870	.993	.993	.997	.997
.5	.700	--	.949	--	.974	--
.7	.232	--	.747	--	.881	--
.9	--	--	--	--	--	--

*The figures given represent percentages of the sum M_o.Where no figures are given, $M_o = T_o = 0$.

i.e., covariability of the r_j's exist, these covariances in many cases, do not contribute significantly to the approximate power of the max-χ^2 test when the sample size n is large. However, we note that the simplified approximation given in (4.2.5) will tend to overestimate the true power of the max-χ^2 test to a slightly greater degree than the full Taylor-series approximation.

4.3 Summary

We have presented an extensive simulation study of the powers of three tests designed to detect serial correlation of the errors in regression in the presence of low order autoregressive models: the Box-Pierce test, the max-χ^2 test and the Durbin-Watson d test. Also, the accuracy of the power approximations of the Box-Pierce and max-χ^2 tests was investigated.

As expected, the Durbin-Watson d test performs extremely well when the regression residuals follow a first order (lag 1) autoregressive process. Since the AR(1) process is a very important case in practice, we conclude that it is always worthwhile to apply the d test. However, our simulation results show that this procedure attains a very low power when the autoregressive model is not first order of lag 1. The d test, then, should be supplemented by another test which has high power in this non-first order case. Our simulation results indicate that, of the two candidates for the supplementary test, i.e., the Box-Pierce and max-χ^2 tests, the max-χ^2 test consistently outperformed the Box-Pierce procedure. The differences in the powers of the two tests was most noticeable when the lag parameter of the alternative error model was negative and when the maximum lag of the model (J) was small compared to the number of sample autocorrelations included in the test statistics (K). We recommend, then, the max-χ^2 test as the most logical alternative to the Durbin-Watson d test, assuming a low-order (small number of non-zero lags) autoregressive model.

Our investigation of the three approximations to the Box-Pierce power--Imhoff's numerical integration, Rao's asymptotic normal square root transformation and Imhoff's extension to Pearson's 3-moment approximation--indicates that, although all three approximations work fairly well for certain values of the lag parameter, the numerical integration technique is the overall most accurate procedure. When the sample size is large (n=200), the numerical integration technique gives an excellent approximation to the asymptotic power of the Box-Pierce test.

The Taylor-series approximation to the power of the max-χ^2 test performs reasonably well for large n, but tends to consistently overestimate the asymptotic power of the test. The results of the study indicate that the contribution of first term of the series expansion to the overall approximate power expression is nearly 100% of the total sum of all the terms in the series for most cases. Thus, the researcher may simplify calculations considerably by considering only the first term in the Taylor-series as an estimate of the power of the max-χ^2 test for large n.

CHAPTER V

OTHER ASYMPTOTIC PROPERTIES OF THE BOX-PIERCE AND MAX-χ^2 TESTS

We have seen (Section 3.3) that the exact distributions of the Box-Pierce and max-χ^2 test statistics become intractible even in those cases when the alternative error model is close to white noise (i.e., a model with a single nonzero lag parameter). In order to compare the two competing tests, we turned to their asymptotic properties. In Chapter 4 the powers of the tests were simulated using critical regions determined from the respective asymptotic null distributions of the test statistics, for sample sizes of 50 and 200. Our goal in Chapter 5 is to investigate the rate at which the powers of the tests tend to one as the sample size increases. Particularly, we wish to determine the test with the faster rate of convergence under various alternatives specified by J (the model lag) and K (the number of sample autocorrelations included in the test statistics).

5.1 Asymptotic Powers for Large n

In this section, the asymptotic powers of the Box-Pierce and max-χ^2 tests are computed for sample sizes ranging from 50 to 5000. The best and most reliable method of generating the asymptotic powers of the tests is to employ the simulation technique outlined in Section 4.1. Using N=1000 iterations, a most conservative bound on the error of estimation of the power is

$$2\sqrt{\frac{(.5)(.5)}{1000}} = .032.$$

However, this technique, even for samples as small as n=50, requires an extraordinary amount of computer time. Since the asymptotic distributions of the test statistics provide good approximations to the true distributions for very large n, we resort to those approximate methods of computing the asymptotic power discussed in Sections 3.3 and 4.2. Recall that the method which best approximates the asymptotic power of the Box-Pierce test is Imhoff's numerical integration technique. The method is only an approximation in the sense that the integral which appears in the power expression can only be evaluated numerically. Based on the power results for n=200, Imhoff's method should provide an excellent approximation to the asymptotic Box-Pierce power for the large values of n which we consider. Our only approximation to the asymptotic max-χ^2 power--the Taylor-series expansion technique--improved greatly when the sample size was increased from 50 to 200. One would expect the method to improve even more for n larger than 200. However, the expansions become unwieldy for K as small as 4 (Sections 3.3.3.2, 4.2.2 and 4.2.3). For purposes of this section, we only consider the initial term in the Taylor-series expansion of the max-χ^2 asymptotic power. From our investigation in Section 4.2.5, the relative contribution of this first term to the overall power is, in many cases, nearly 100%, thus providing only a slight overestimate of the asymptotic power of the max-χ^2 test.

The alternative error model under which the powers of the tests are computed is the first order AR lag J model

$$Z_t = \emptyset Z_{t-J} + \varepsilon_t, \tag{5.1.1}$$

where $\{\varepsilon_t\}$ is a white noise process. The various values of J, K, ϕ and n considered in this study are given in the following table.

As stated in the introduction to this chapter, we are particularly interested in determining which asymptotic power converges to one at the fastest rate for a given value of K. Our initial plan was to consider K=4, 12 and 24, consistent with the values of K employed in the simulation study. However for K=12 and K=24 we experienced great difficulties in computing the value of the integral in Imhoff's Box-Pierce power expression. The problem is that the variance-covariance matrix (V) of the asymptotic distribution of the first 12 (or 24) sample autocorrelations is very nearly singular (e.g., $|V| = 7\text{x}10^{-6}$ for J=2, K=24). These near singularities prevent the numerical integration technique that we used (available in the IMSL FORTRAN package) from computing reliable values in some cases, and in most cases, any values at all. Unable to obtain asymptotic powers of the Box-Pierce test for K=12 and K=24, we added the case K=8 in order to compare rates of convergence across at least two different values of K.

We present a portion of the results in the form of figures[1], as in Chapter 4. Once again, the complete set of numerical results is given in the Appendix (Tables A-41 - A-44).

Figures 5.1 - 5.4 represent power curves for the tests as n is increased to infinity and the values of J, K and ϕ are kept fixed. It is interesting to note that for very small values of the lag parameter (ϕ = -.1 or .1), a sample size of at least 2000 is needed before the tests

---

[1] Note that the horizontal axis in Figures 5.1 - 5.4 is not drawn to scale.

TABLE 5.1

ASYMPTOTIC POWER COMPARISONS
BOX-PIERCE AND MAX-χ^2 TESTS
ERROR MODEL: $Z_t = \emptyset Z_{t-5} + \varepsilon_t$

Parameters:	(J,K)	$\emptyset$	n	α
	(1,4)	-.5	50	.05
	(2,4)	-.3	100	
	(3,4)	-.1	200	
	(4,4)	.1	500	
	(1,8)	.3	1000	
	(2,8)	.5	2000	
	(3,8)		5000	

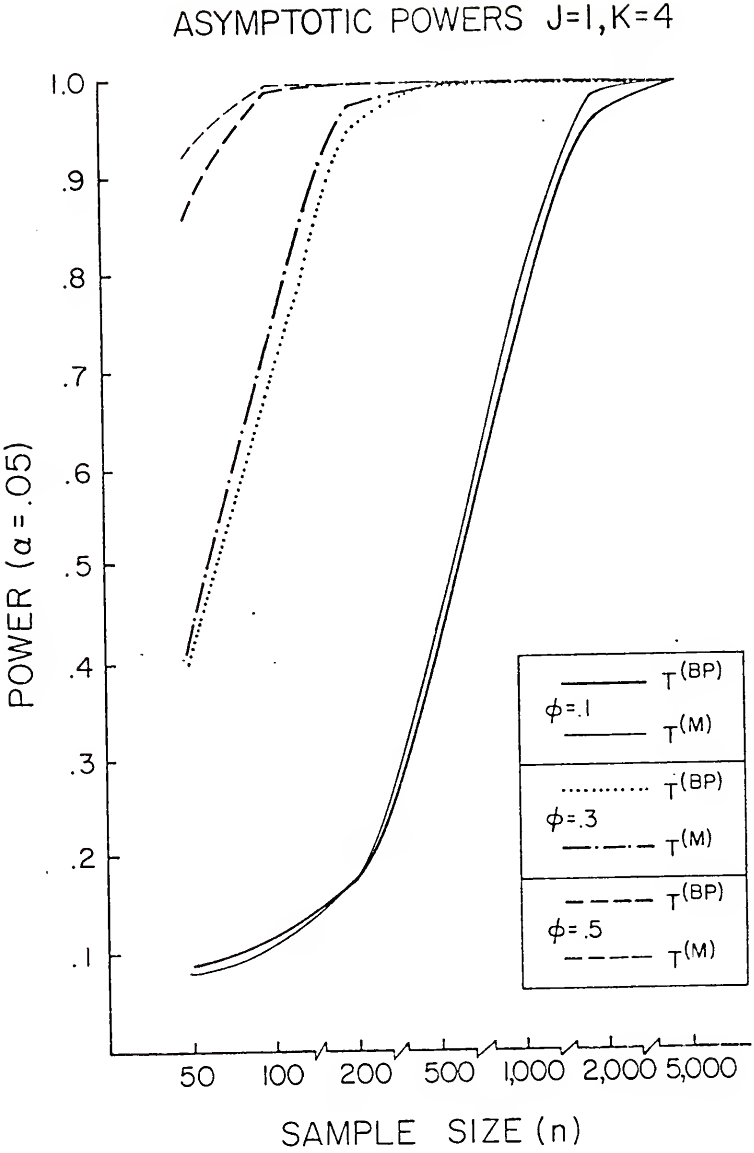

ASYMPTOTIC POWERS J=1, K=4
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
50
100
200
500
1,000
2,000
5,000
SAMPLE SIZE (n)
ϕ=.1
$T^{(BP)}$
$T^{(M)}$
ϕ=.3
$T^{(BP)}$
$T^{(M)}$
ϕ=.5
$T^{(BP)}$
$T^{(M)}$

FIGURE 5.1

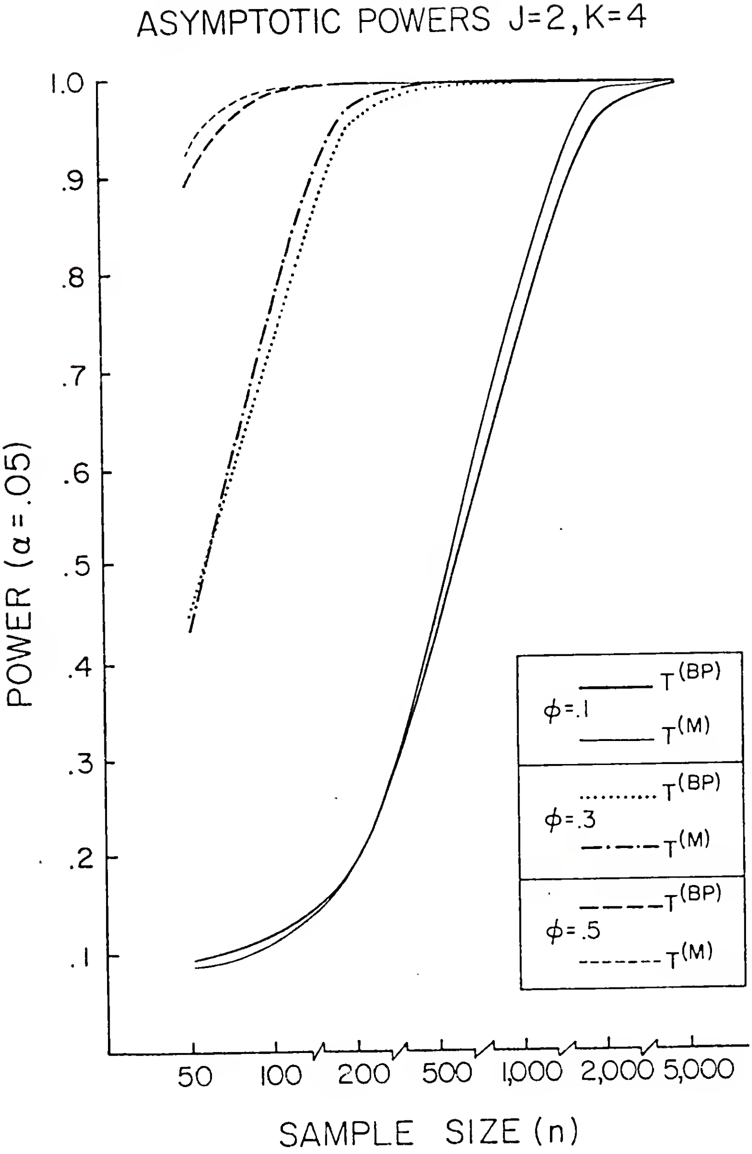

ASYMPTOTIC POWERS J=2, K=4
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
50
100
200
500
1,000
2,000
5,000
SAMPLE SIZE (n)
φ=.1
T(BP)
T(M)
φ=.3
T(BP)
T(M)
φ=.5
T(BP)
T(M)

FIGURE 5.2

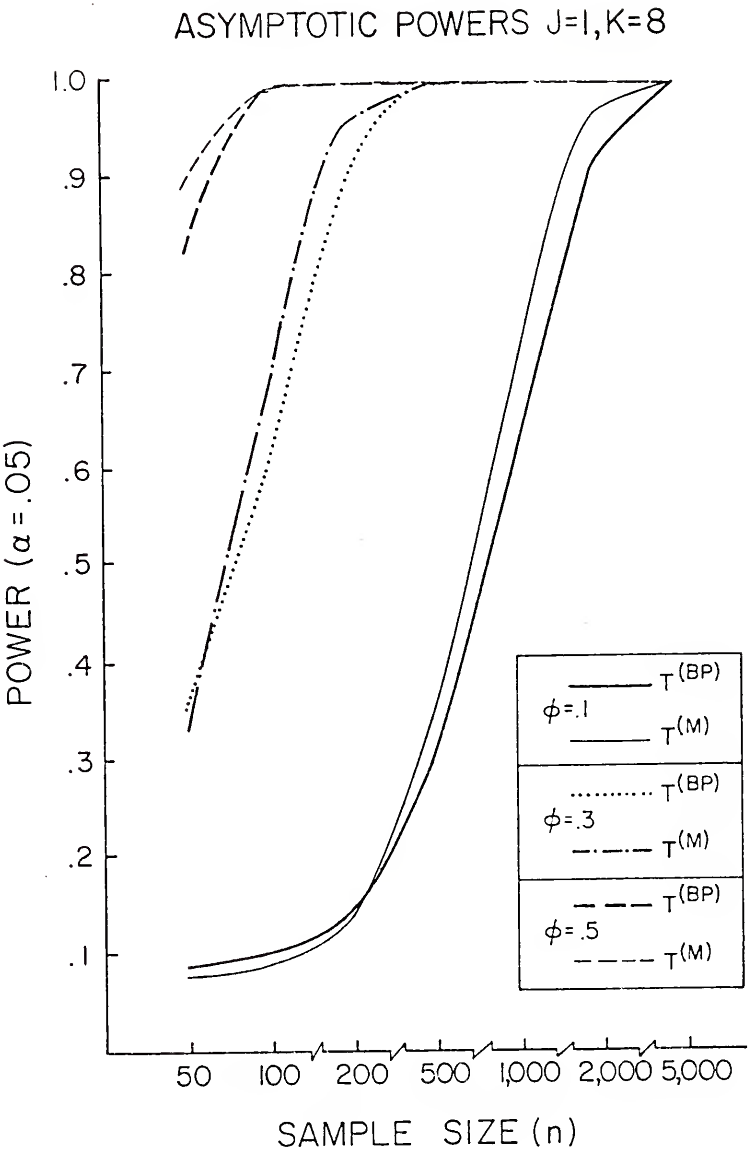

ASYMPTOTIC POWERS J=1, K=8
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
50
100
200
500
1,000
2,000
5,000
SAMPLE SIZE (n)
ϕ=.1
$T^{(BP)}$
$T^{(M)}$
ϕ=.3
$T^{(BP)}$
$T^{(M)}$
ϕ=.5
$T^{(BP)}$
$T^{(M)}$

FIGURE 5.3

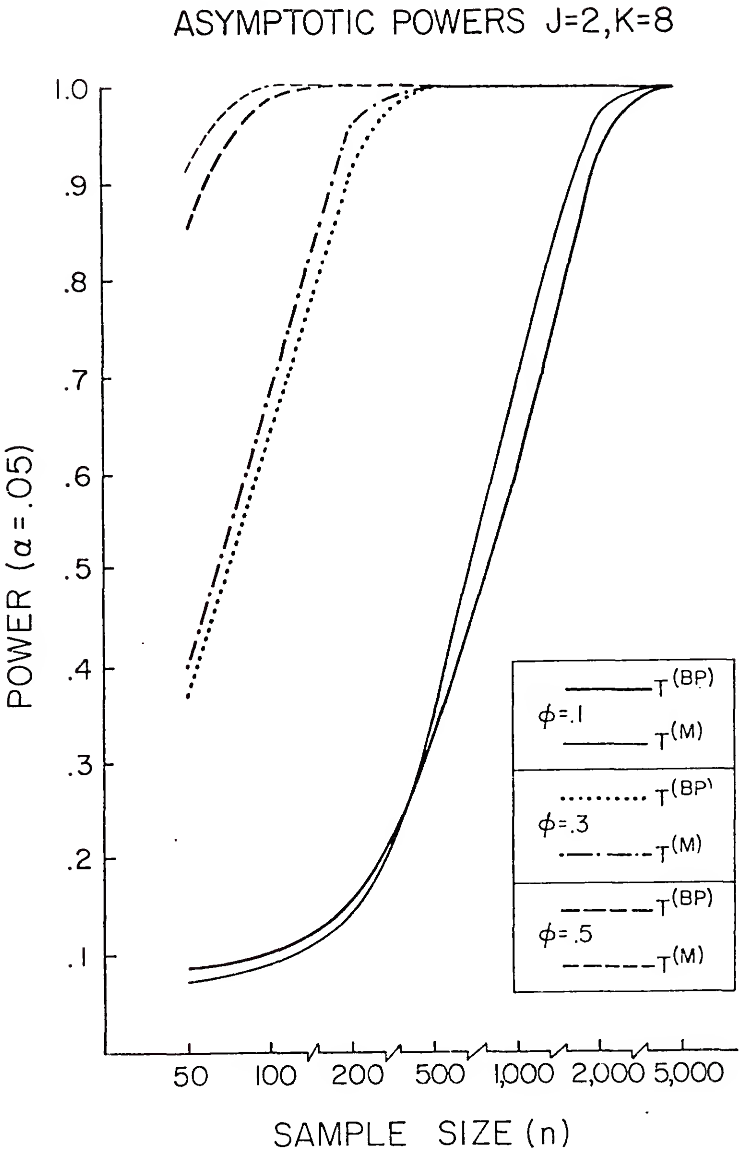

ASYMPTOTIC POWERS J=2,K=8
POWER (α = .05)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
50
100
200
500
1,000
2,000
5,000
SAMPLE SIZE (n)
φ=.1
T(BP)
T(M)
φ=.3
T(BP)
T(M)
φ=.5
T(BP)
T(M)

FIGURE 5.4

attain an asymptotic power of 1, while for moderate values of ϕ (-.5 or .5), powers of 1 would be attained for n=100 if the true distribution of the test statistic had converged to the asymptotic distribution. A comparison of the asymptotic powers of the tests reveals that, for each combination of J, K and ϕ shown in the figures, the asymptotic power of the max-χ^2 test converges to 1 at a faster rate than the asymptotic power of the Box-Pierce test, i.e., the minimum value of n needed for the max-χ^2 test to attain a power of 1 is smaller than the corresponding value for the Box-Pierce test. This result also holds true for those cases not presented as figures but given in the Appendix. In particular, note the results for $|\phi| = .1$. At this low value of the lag parameter, the Box-Pierce test dominates the max-χ^2 test when n is 200 or smaller. This result agrees with our simulations in Chapter 4. However, for n between 200 and 500, the power curves cross and the max-χ^2 test begins to dominate. From earlier results, the researcher who attempts to detect serial correlation of the errors in the presence of error models of the form of (5.1.1) with low lag coefficients using either the Box-Pierce or max-χ^2 test and sample sizes of 50 or 200, will fail to reject the null hypothesis of white noise at a very high rate. Neither of the two tests provide the researcher with a reliable procedure. But when n is increased to 1000, both tests attain reasonable power, and the max-χ^2 test is the more powerful test.

From simulation at n=50 and n=200, we noted that the max-χ^2 procedure appears to be less sensitive to an increase in K, in terms of power loss. This result also holds true for very large values of n. Compare Figures 5.2 and 5.4, the case (J=2, K=4) versus the case (J=2, K=8). Especially at the low lag parameter value of .1, the difference between the max-χ^2 power and Box-Pierce power noticeably increases (for n greater than 500) as K is increased from 4 to 8. Simulation at n=50 and n=200 indicates that this difference is even greater for K=12 and K=24. Thus, the max-χ^2 test attains a power of 1 with a significantly smaller sample size than is needed for the Box-Pierce test.

From the results presented in this section, we conclude that for alternative error models given by (5.1.1) with J (the model lag), $\emptyset$ (the lag parameter) and K (the number of autocorrelations included in the test statistic) fixed, the max-χ^2 asymptotic power converges to 1 at a faster rate than the Box-Pierce asymptotic power as n (the sample size) tends to infinity. In Section 5.2, we investigate a type of asymptotic relative efficiency, due to Hodges and Lehmann, which relates closely to the problem of power convergence.

5.2 Hodges-Lehmann Asymptotic Relative Efficiency

J. L. Hodges, Jr. and E. L. Lehmann [21] developed an alternative notion of asymptotic efficiency, designed to measure the "speed with which power at a fixed alternative tends to 1." The relevance of this type of efficiency to the results of the preceding section is clear. Bascially, the objective of Hodges and Lehmann asymptotic relative efficiency (HLARE) is to compare the rates at which the power functions of the tests converge to 1, for a fixed alternative (θ) and fixed significance level (α). The test with the faster rate of convergence is, of course, deemed superior. A more rigorous definition follows.

Definition 3 (Hodges-Lehmann): Define $\beta_A(N,\alpha,\theta)$ to be the power of test A at significance level α under the alternative θ, using N observations. Define $\beta_A*(N*,\alpha,\theta)$ for test A* similarly. Then the relative efficiency of A* to A is N/N*, where N* is the sample size needed for test A* to have power $\beta_A(N,\alpha,\theta)$, i.e., $\beta_A*(N*,\alpha,\theta) = \beta_A(N,\alpha,\theta)$. □

Hodges and Lehmann (H-L) note that there usually will not exist an integer N* such that $\beta_A^*(N^*,\alpha,\theta) = \beta_A(N,\alpha, \theta)$, but rather an integer N_o such that $\beta_A^*(N_o,\alpha,\theta) < \beta_A(N,\alpha,\theta) < \beta_A^*(N_o+1,\alpha,\theta)$. Given the existence of N_o, the H-L technique for choosing N* is stated as follows:

Definition 4: Define N* of Definition 3 to be N_o+p, where $\beta_A^*(N_o,\alpha,\theta) < \beta_A(N,\alpha,\theta) < \beta_A^*(N_o+1,\alpha,\theta)$ and where the test A* has power $\beta_A(N,\alpha,\theta)$ if its number of observations is randomly chosen as (N_o+1) with probability p and N_o with probability 1-p. ☐

The value of N* defined above, then, is simply the expected number of observations required for test A* to match the power of test A, when randomizing between consecutive integers. Given this value of N*, the HLARE is defined as follows.

Definition 5: The Hodges-Lehmann asymptotic relative efficiency (HLARE) of test A* to A is defined to be

$$\lim_{N\to\infty} \frac{N}{N^*} ,$$

where N and N* are defined as in Definitions 3 and 4. ☐

H-L also give a technique useful for calculating the HLARE. The result is stated here as a lemma.

Lemma 15 (Hodges-Lehmann): Define $\beta(\theta) = \beta_A(N,\alpha,\theta)$ to be the power of test A at significance level α under the alternative θ, using N observatons, and $\beta^*(\theta) = \beta_A^*(N,\alpha,\theta)$ similarly. Suppose also that

$$\sqrt[N]{1-\beta(\theta)} \longrightarrow b(\theta) \quad \text{as} \quad N \to \infty$$

and

$$\sqrt[N]{1-\beta^*(\theta)} \longrightarrow b^*(\theta) \quad \text{as} \quad N \to \infty \ .$$

If N* is defined as in Definition 4, then

$$\frac{N^*}{N} \longrightarrow \frac{\ln b^*(\theta)}{\ln b(\theta)} \quad \text{as } N \to \infty .$$

Proof: See Hodges-Lehmann [21], ppg. 329-333. □

As a result of Lemma 15, the HLARE of A to A* may be obtained by computing the ratio

$$e_{A,A^*} = \frac{\ln b^*(\theta)}{\ln b(\theta)} \quad .$$

Ideally, we would like to compute the HLARE of $T^{(BP)}$ to $T^{(m)}$, where $T^{(BP)}$ and $T^{(m)}$ are the Box-Pierce and max-χ^2 tests, respectively. However, two main problems are encountered. First, since the exact non-null distributions of the test statistics are unknown, the exact power function of each test is intractible and can only be approximated using the asymptotic distributions. Hodges-Lehmann make no mention of the possible drawbacks of using asymptotic power functions in place of exact power functions in HLARE theory. Secondly, given that we assume that the asymptotic power expressions closely approximate the exact powers, the limit of $\sqrt[N]{1-\beta(\theta)}$ as $N \to \infty$ is not easily computed. In the case of the max-χ^2 statistic, the power function is approximated by a Taylor-series expansion of the asymptotic density function of the first K sample autocorrelations. The terms of the infinite expansion may or may not be functions of N, depending upon the structure of the alternative error model. It is difficult to determine what effect those terms omitted from the expansion will have on HLARE. For the Box-Pierce test, the problem is even greater. Imhoff's approximation to the Box-Pierce asymptotic power involves an integral which cannot be written in closed form. Our only recourse is to use either of the other two Box-Pierce power approximations described in Section 3.3. Of these, though, Imhoff's extension to Pearson's

3-moment approximation also becomes unmanageable, since the method involves evaluating a tail probability of a non-central χ^2 random variable. Rao's asymptotic square root transformation, however, permits a workable approximate power function. Although the application of the HLARE procedure to the test statistics using approximate asymptotic powers cannot guarantee reliable results, we feel that an attempt at obtaining the "approximate" HLARE, will at the very least, provide valuable insight into the power convergence question.

We consider testing the hypotheses H_o: $\phi = 0$ versus H_a: $\phi \neq 0$, under the alternative error model

$$Z_t = \phi Z_{t-4} + \varepsilon_t \quad , \tag{5.2.1}$$

where the $\{\varepsilon_t\}$ is a white noise process. The number of autocorrelations that we choose to include in our test statistics is four, hence the case J=4, K=4. We select this rather simple case in order to eliminate the problem of omitting terms of order n in the Taylor-series expansion of the max-χ^2 power. Equation (4.2.2) from the Theorem 12 shows that n appears in only the first term of the series expansion. In general,this will be true as long as J=K. Thus, although the HLARE of $T^{(BP)}$ to $T^{(m)}$ is stated and proved for J=K=4, the result applies to the more general case J=K. The proof, however, requires a lemma which we state first.

Lemma 16: For large N, the expression $\Phi(K-\sqrt{N}\delta)$ is equivalent (in the sense of ratio) to

$$\left(\frac{1}{\sqrt{N}\delta}\right)\left(\frac{1}{\sqrt{2\pi}}\right) \exp\{-(\tfrac{1}{2})(\sqrt{N}\delta-K)^2\} \quad , \tag{5.2.2}$$

where

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2} dz \quad ,$$

K and δ are fixed.

Proof: From [3] pg. 932, an approximation for $1-\Phi(x)$ for $x > 0$

$$\frac{\frac{1}{\sqrt{2\pi}}\exp\{-\frac{1}{2}x^2\}}{x}\left(1-\frac{a_1}{x^2+2}+\frac{a_2}{(x^2+2)(x^2+4)}+\cdots\right).$$

where a_1, a_2, a_3, . . . are constants which do not depend on x. Since $1-\Phi(x) = \Phi(-x)$, we have for $x < 0$

$$\Phi(x) \sim \frac{\frac{1}{\sqrt{2\pi}}\exp\{-\frac{1}{2}x^2\}}{-x}\left(1-\frac{a_1}{x^2+2}+\frac{a_2}{(x^2+2)(x^2+4)}+\cdots\right) , \tag{5.2.3}$$

where ~ means "is approximate equal to." Substituting $K-\sqrt{N}\delta$ for x in equation (5.2.3), we obtain

$$\Phi(K-\sqrt{N}\delta) \sim \frac{\frac{1}{\sqrt{2\pi}}\exp\{-\frac{1}{2}(\sqrt{N}\delta-K)^2\}}{\sqrt{N}\delta-K}\left[1-\frac{a_1}{(K-\sqrt{N}\delta)^2+2}+\cdots\right].$$

For large N then, we have

$$\Phi(K-\sqrt{N}\delta) \sim \frac{\frac{1}{\sqrt{2\pi}}\exp\{-\frac{1}{2}(\sqrt{N}\delta-K)^2\}}{\sqrt{N}\delta} ,$$

which is the desired result. □

Theorem 13: The HLARE of $T^{(BP)}$ to $T^{(m)}$ for the case (J=4, K=4) is 1.

Proof: This proof utilizes the result given by Lemma 16. First consider the max-χ^2 test, $T^{(m)}$. From Equation (4.2.2), the power (asymptotic) of $T^{(m)}$ can be expressed as

$$\beta(N,\alpha,\emptyset) = 1-[\Phi(K-\sqrt{N}\delta)-\Phi(-K-\sqrt{N}\delta)][A] \tag{5.2.4}$$

where

$$K = \frac{\sqrt{C_\alpha^{(BP)}}}{\sqrt{1-\emptyset^2}} , \quad \delta = \frac{\emptyset}{\sqrt{1-\emptyset^2}}$$

and A is a constant free of N. Then,

$$\sqrt[N]{1-\beta(N,\alpha,\emptyset)} = [\Phi(K-\sqrt{N}\delta) - \Phi(-K-\sqrt{N}\delta)]^{\frac{1}{N}} \{A\}^{\frac{1}{N}} \tag{5.2.5}$$

Utilizing the result from Lemma 17, we may write the right-hand side of equation (5.2.5) as

$$(\frac{1}{\sqrt{N}\delta})^{\frac{1}{N}}(\frac{1}{\sqrt{2\pi}})^{\frac{1}{N}} [\exp\{-\frac{1}{2}(K-\sqrt{N}\delta)^2\}-\exp\{-\frac{1}{2}(-K-\sqrt{N}\delta)^2\}]^{\frac{1}{N}} \{A\}^{\frac{1}{N}}$$

$$= (\frac{1}{\delta\sqrt{2\pi}})^{\frac{1}{N}}(N)^{-\frac{1}{2N}}[e^{-\frac{1}{2}N\delta^2}]^{\frac{1}{N}}[e^{-\frac{1}{2}K^2}]^{\frac{1}{N}}[e^{\sqrt{N}\delta K}-e^{-\sqrt{N}\delta K}]^{\frac{1}{N}} \{A\}^{\frac{1}{N}}$$

$$= (\frac{1}{\delta\sqrt{2\pi}})^{\frac{1}{N}}(e^{-\frac{1}{2N}\ln N})[e^{-\frac{1}{2}\delta^2}][e^{-\frac{1}{2}K^2}]^{\frac{1}{N}}\exp\{\frac{1}{N}\ln(^{\sqrt{N}\delta K}-e^{-\sqrt{N}\delta K})\}\{A\}^{\frac{1}{N}}$$

$$= (\frac{1}{\delta\sqrt{2\pi}})^{\frac{1}{N}}(e^{-\frac{1}{2N}\ln N})[e^{-\frac{1}{2}\delta^2}][e^{-\frac{1}{2}K^2}]^{\frac{1}{N}} \exp\{\frac{1}{N}\ln(e^{\sqrt{N}\delta K})(1+0(1))\}\{A\}^{\frac{1}{N}}$$

$$= (\frac{1}{\delta\sqrt{2\pi}})^{\frac{1}{N}}(e^{-\frac{1}{2N}\ln N})[e^{-\frac{1}{2}\delta^2}][e^{-\frac{1}{2}K^2}]^{\frac{1}{N}}\exp\{\frac{\delta K}{\sqrt{N}} (1+0(1))\} \{A\}^{\frac{1}{N}}$$

In this form, it is easy to see that

$$\sqrt[N]{1-\beta(N,\alpha,\emptyset)} \longrightarrow \exp\{-\frac{1}{2}\delta^2\} \quad \text{as } N \to \infty$$

where

$$\delta = \frac{\emptyset}{\sqrt{1-\emptyset^2}} \quad .$$

Thus, for $T^{(m)}$,

$$b(\emptyset) = \exp\{-(\frac{1}{2})\frac{\emptyset^2}{(1-\emptyset^2)}\} \quad .$$

Now consider the power expression for the Box-Pierce test, $T^{(BP)}$. Recall from equation 3.3.20, the Box-Pierce power expression using

Rao's square root transformation:

$$\beta^*(N,\alpha,\emptyset) = 1 - \Phi(K^*-\sqrt{N}\delta^*) + \Phi(-K^*-\sqrt{N}\delta^*) \qquad (5.2.6)$$

where

$$K^* = \frac{\sqrt{C_\alpha^{(BP)}}}{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^{-\frac{1}{2}}(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}} \quad , \quad \delta^* = \frac{(\underset{\sim}{\rho}'\underset{\sim}{\rho})}{(\underset{\sim}{\rho}'V\underset{\sim}{\rho})^{\frac{1}{2}}}$$

and $\underset{\sim}{\rho}$ and V are the mean vector and covariance matrix, respectively, of the asymptotic distribution of the first K sample autocorrelations. The result that

$$\sqrt[N]{1-\beta^*(N,\alpha,\emptyset)} \longrightarrow \exp\{-\frac{1}{2}(\delta^*)^2\} \quad \text{as } N \to \infty \quad ,$$

follows directly from the max-χ^2 case. We are left, then, to determine (δ^*) for the case (J=4, K=4). From Section 4.2.3, we have, for this case:

$$\underset{\sim}{\rho} = (0,\ 0,\ 0,\ \emptyset)' \ ,$$

and

$$V = \begin{bmatrix} \frac{1+\emptyset^2}{1-\emptyset^2} & 0 & \frac{2\emptyset}{1-\emptyset^2} & 0 \\ 0 & \frac{(1+\emptyset)^2}{1-\emptyset} & 0 & 0 \\ \frac{2\emptyset}{1-\emptyset^2} & 0 & \frac{1+\emptyset^2}{1-\emptyset^2} & 0 \\ 0 & 0 & 0 & 1-\emptyset^2 \end{bmatrix}$$

Thus, $\underset{\sim}{\rho}'\underset{\sim}{\rho} = \emptyset^2$ and $\underset{\sim}{\rho}'V\underset{\sim}{\rho} = \emptyset^2(1-\emptyset^2)$. Hence,

$$(\delta^*)^2 = \frac{(\underset{\sim}{\rho}'\underset{\sim}{\rho})^2}{\underset{\sim}{\rho}'V\underset{\sim}{\rho}} = \frac{\emptyset^4}{\emptyset^2(1-\emptyset^2)} = \frac{\emptyset^2}{1-\emptyset^2} \quad ,$$

and for $T^{(BP)}$, $b^*(\emptyset) = \exp\{-(\frac{1}{2})\frac{\emptyset^2}{(1-\emptyset^2)}\}$. Since $b(\emptyset)$ and $b^*(\emptyset)$ are equivalent, we have, in the case (J=4, K=4), HLARE of $T^{(BP)}$ to $T^{(m)}$ equal to 1. ☐

At first glance, the result of Theorem 13 seems to contradict the asymptotic power comparisons for large n of Section 5.1, where we concluded that, for fixed J, K, and $\emptyset$, the minimum sample size needed (n) for the max-χ^2 power to equal 1 was less than the corresponding sample size (n*) for the Box-Pierce test. However, note that the slopes of the power curves in Figure 5.1 - 5.4 are nearly indentical, and although the max-χ^2 does attain a power of 1 first, the difference n*-n does not appear to be very large. We would expect, then, $\frac{n^*}{n} \longrightarrow 1$ as n infinitely increases. Since the HLARE is essentially the limiting ratio of the sample sizes needed for the tests to attain equal powers for fixed $\emptyset$, α, J, K, the result is not surprising. However, Hodges and Lehmann [22] state that if two tests have HLARE = 1, this does not necessarily imply that the tests are asymptotically equivalent. In fact, they point out that when $\text{HLARE} = \lim_{n\to\infty} \frac{n^*}{n} = 1$, it may be more meaningful to consider the more natural quantity, $\text{HLAD} = \lim_{n\to\infty}(n^*-n)$, called the asymptotic deficiency. If the limit exists, HLAD can be interpreted as a measure of the additional number of observations needed by $T^{(BP)}$ to attain the same power as $T^{(m)}$ for fixed K, J, $\emptyset$ and α. Considering the asymptotic power results of Section 5.1, this concept of asymptotic deficiency is worthy of further investigation.

The basic technique outlined by Hodges-Lehmann [22] is to equate the powers of the two competing tests, with respective sample sizes of n and n*. Letting $d_n = n^*-n$, it may be possible to show that d_n converges as n tends to infinity. We compute the HLAD of $T^{(BP)}$ to $T^{(m)}$, where we take n* to be the sample size associated with $T^{(BP)}$ and n to be the sample size associated with $T^{(m)}$. Again, the result is given in the following theorem for J=K=4, but also applies to the more general J=K case.

Theorem 14: Let $\beta(N,\alpha,\emptyset)$ be the power of the max-χ^2 procedure under the alternative error model given by (5.2.1). Let N^* be the required sample size for the Box-Pierce test to have equivalent power, and define $d_N = N^*-N$. Then for the case $J=K=4$,

$$\lim_{N\to\infty} d_N = \infty$$

Proof: From Theorem 13, equation (5.2.4), the power of the max-χ^2 test can be expressed as

$$\beta(N,\alpha,\emptyset) = 1-[\Phi(K_1-\sqrt{N}\delta) - \Phi(-K_1-\sqrt{N}\delta)][A] ,$$

where

$$K_1 = \frac{\sqrt{C_\alpha^{(m)}}}{\sqrt{1-\emptyset^2}} \quad , \quad \delta = \frac{\emptyset}{\sqrt{1-\emptyset^2}}$$

and A a constant. Similarly, from equation (5.2.6), the Box-Pierce power is given as

$$\beta^*(N^*,\alpha,\emptyset) = 1-[\Phi(K_2-\sqrt{N^*}\delta) - \Phi(-K_2-\sqrt{N^*}\delta)] ,$$

where

$$K_2 = \frac{\sqrt{C\alpha^{(BP)}}}{\sqrt{1-\emptyset^2}}$$

and δ as above. In order to find the HLAD, we need to equate the two powers. First, consider the following. Assume $\delta > 0$ (i.e., $\emptyset$ positive). Then for large N

$$\Phi(K-\sqrt{N}\delta) - \Phi(-K-\sqrt{N}\delta) \approx \Phi(K-\sqrt{N}\delta),$$

in the sense of ratio, since $\lim \frac{\Phi(-K-\sqrt{N}\delta)}{\Phi(K-\sqrt{N}\delta)} = 0$ (L`Hospital).

Conversely, for $\delta>0$ (i.e., $\emptyset$ negative), $\Phi(K-\sqrt{N}\delta)-\Phi(-K-\sqrt{N}\delta) \approx 1-\Phi(-K-\sqrt{N}\delta)$

when N is large. We use these results to simplify the power functions given above.

First, consider the case $\delta > 0$. Then $\beta(N,\alpha,\emptyset) \approx \Phi(K_1-\sqrt{N}\delta)(A)$ and $\beta^*(N^*,\alpha,\emptyset) \approx \Phi(K_2-\sqrt{N^*}\delta)$. Equating the powers of the tests, we write

$$[\Phi(K_1-\sqrt{N}\delta)][A] = \Phi(K_2-\sqrt{N^*}\delta).$$

Applying Lemma 17, we obtain

$$\left(\frac{1}{\sqrt{N}\delta}\right)\left(\frac{1}{\sqrt{2\pi}}\right)\exp\{-\frac{1}{2}(K_1-\sqrt{N}\delta)^2\}\cdot A = \left(\frac{1}{\sqrt{N^*}\delta}\right)\left(\frac{1}{\sqrt{2\pi}}\right)\exp\{-\frac{1}{2}(K_2-\sqrt{N^*}\delta)^2\}$$

$$\Longleftrightarrow \quad \left(\frac{1}{\sqrt{N}\delta}\right)\exp\{-\frac{1}{2}(K_1^2-2\sqrt{N}K_1\delta+N\delta^2)\}\ A=\left(\frac{1}{\sqrt{N^*}\delta}\right)\exp\{-\frac{1}{2}(K_2^2-2\sqrt{N^*}K_2\delta+N^*\delta^2)\}$$

$$\Longleftrightarrow \quad \left(\frac{1}{\sqrt{N}\delta}\right)e^{-\frac{1}{2}K_1^2}\cdot e^{\sqrt{N}K_1\delta}\cdot e^{-\frac{N}{2}\delta^2}(A) = \left(\frac{1}{\sqrt{N^*}\delta}\right)e^{-\frac{1}{2}K_2^2}\cdot e^{\sqrt{N^*}K_2\delta}\cdot e^{-\frac{N^*}{2}\delta^2}$$

$$\Longleftrightarrow \quad \sqrt{\frac{N^*}{N}}\, e^{\frac{1}{2}(K_2^2-K_1^2)}(A) = e^{-\frac{1}{2}\delta^2(N^*-N)}\cdot e^{(K_2\delta\sqrt{N^*}-K_1\delta\sqrt{N})} \quad . \qquad (5.2.7)$$

Now let $d_N=N^*-N$ and consider the last term on the right-hand side of equation (5.2.7). Substituting d_N+N for N^* and expanding, we have

$$\begin{aligned} K_2\delta\sqrt{N+d_N}-K_1\delta\sqrt{N} &= \delta[K_2\sqrt{N+d_N}-K_1\sqrt{N}] \\ &= \sqrt{N}\delta[K_2\sqrt{1+\frac{d_N}{N}}-K_1] \\ &= \sqrt{N}\delta[K_2\{1+\frac{d_N}{2N}+0(\frac{1}{N})\}-K_1] \quad . \end{aligned}$$

Substituting back into equation (5.2.7), we obtain

$$\sqrt{1+\frac{d_N}{N}}\ e^{\frac{1}{2}(K_2^2-K_1^2)}(A) = e^{-\frac{1}{2}\delta^2(d_N)}\cdot e^{\sqrt{N}\delta[K_2\sqrt{1+\frac{d_N}{N}}-K_1]}$$

$$\Longleftrightarrow \quad \left(1+\frac{d_N}{N}\right)e^{(K_2^2-K_1^2)}(A^2) = e^{-d_N\delta^2}\cdot e^{2\sqrt{N}\delta\,[K_2\sqrt{1+\frac{d_N}{N}}-K_1]}$$

Taking logs of both sides, gives

$$2\ln(A) + \ln(1+\frac{d_N}{N}) + (K_2^2-K_1^2) = -\delta^2(d_N)+2\sqrt{N}\delta[K_2\sqrt{1+\frac{d_N}{N}} -K_1]$$

$$\Longleftrightarrow \quad \delta^2(d_N) = -2\ln(A)-\ln(1+\frac{d_N}{N})+(K_1^2-K_2^2) +$$

$$2\sqrt{N}\delta[K_2\{1+\frac{d_N}{2N} + 0(\frac{1}{N})\} -K_1]$$

$$\Longleftrightarrow \quad d_N = \frac{-2\ln(A)}{\delta^2} - \frac{\ln(1+\frac{d_N}{N})}{\delta^2} + \frac{(K_1-K_2^2)}{\delta^2} + \frac{2\sqrt{N}}{\delta}[K_2\{1+\frac{d_N}{2N} + 0(\frac{1}{N})\}-K_1]$$

Now, since $\lim_{N\to\infty} \frac{N^*}{N} = 1$ (Theorem 13), then $\lim_{N\to\infty} \frac{d_N}{N} = 0$. Thus,

$$\lim_{N\to\infty} d_N = \begin{cases} \infty \text{ if } \frac{K_2-K_1}{\delta} > 0 \\ -\infty \text{ if } \frac{K_2-K_1}{\delta} < 0 \end{cases} .$$

Since $C_\alpha^{(BP)} > C_\alpha^{(m)}$ for all α, then $K_2>K_1$ and $\frac{K_2-K_1}{\delta} > 0$. Thus,

$$\lim_{N\to\infty} d_N = \infty.$$

For $\delta<0$, we have $\Phi(K-\sqrt{N}\delta)-\Phi(-K-\sqrt{N}\delta)\approx\Phi(-K-\sqrt{N}\delta)$ for large N. Hence, we only need to replace K_1 by $-K_1$ and K_2 by $-K_2$ in the above proof. This gives

$$\lim_{N\to\infty} d_N = \begin{cases} \infty \text{ if } \frac{K_1-K_2}{\delta} > 0 \\ -\infty \text{ if } \frac{K_1-K_2}{\delta} < 0 \end{cases} .$$

But $K_2>K_1$ and $\delta<0$ implies that $\frac{K_1-K_2}{\delta} > 0$, hence for δ positive or negative

$$\lim_{N\to\infty} d_N = \lim_{N\to\infty}(N^*-N) = \infty . \qquad \square$$

The result in Theorem 14 implies that the Box-Pierce test is asymptotically deficient when compared to the max-χ^2 procedure, i.e., asymptotically, the additional number of observations required for the Box-Pierce procedure to to attain the same power as the max-χ^2 test tends to infinity. This measure of asymptotic deficiency strongly supports our asymptotic power results of Section 5.1 and provides additional evidence that the max-χ^2 test, for large n, outperforms the Box-Pierce test.

Before discussing the next topic in this chapter, we make two comments. First, we remind the reader that n^* and n represent conservative and liberal estimates, respectively, of the sample sizes needed for the tests to attain identical powers, since, in fact, the Rao square root transformation tends to underestimate the asymptotic Box-Pierce power while the first term of the Taylor-series expansion tends to overestimate the asymptotic max-χ^2 power. Second, the fact that $\frac{n^*}{n} \to 1$ while $n^*-n\to\infty$ implies that $n^*-n=0(n^{\varepsilon})$, where $0<\varepsilon<1$. Knowledge of ε will provide the practical researcher with valuable information on the rate at which $n^*-n\to\infty$, or equivalently, the additional number of observations required for the Box-Pierce test to attain a power equivalent to the max-χ^2 test based on n observations. We leave the problem of determining ε as a reasonable and important topic of future research.

The final section of this chapter will be devoted to a consideration of the likelihood ratio statistic for testing serial correlation of the errors under the alternative error model of (5.1.1).

5.3 Likelihood Ratio

Consider testing the hypothesis H_o: $\emptyset = 0$ under the alternative error model

$$Z_t = \emptyset Z_{t-J} + \varepsilon_t, \tag{5.3.1}$$

where $\{\varepsilon_t\}$ is white noise, and $0<|\emptyset|<1$ when H_a is true. The results presented in Chapter 4 and the first two sections of this chapter all indicate that the max-χ^2 test is a more powerful procedure for detecting this model than the Box-Pierce test. In this section, we investigate the likelihood ratio statistic for two purposes: (1) to determine if the statistic could practically be used in a test for serial correlation of the errors of a regression model and (2) to determine what form of test statistic the likelihood ratio suggests.

5.3.1 General Case

We first derive the likelihood ratio test for the general case of J and $\emptyset$ in model (5.3.1) and K, the number of sample autocorrelations included in the test all known but unspecified. The result is given in the following lemma.

Lemma 18: Consider a test of H_o: $\emptyset = 0$ versus H_a: $\emptyset \neq 0$ based upon the sequence of sample autocorrelations $\{r_1, r_2, r_3, \ldots, r_K\}$, where $\emptyset$ in the alternative and J in model (5.3.1) are known but unspecified and where

$$r_j = \sum_{t=j+1}^{n} Z_t Z_{t-j} / \sum_{t=1}^{n} Z_t^2 , \qquad 1 \leq j \leq K.$$

The likelihood ratio test, for large n, is based upon the statistic

$$\lambda = |V|^{\frac{1}{2}} \exp\{-\frac{n}{2} \sum_{j=1}^{K} r_j^2 + \frac{n}{2}(\underset{\sim}{r}-\underset{\sim}{\rho})' V^{-1} (\underset{\sim}{r}-\underset{\sim}{\rho})\} \qquad (5.3.2)$$

where $\underset{\sim}{\rho}' = (\rho_1, \rho_2, \ldots, \rho_K)$ is the vector of true autocorrelations and V is the variance-covariance matrix of the asymptotic distribution of the first K sample autocorrelations.

Proof: Since the exact distribution of the sample autocorrelations is unknown, we must base our likelihood ratio test upon the asymptotic distribution given in Section 3.3.2. Recall from Lemma 1 that

$$\sqrt{n}\,(\underset{\sim}{r}-\underset{\sim}{\rho}) \xrightarrow{D} N_K\{\underset{\sim}{0}, V\} ,$$

or equivalently, for large n,

$$\sqrt{n}\,\underset{\sim}{r} \stackrel{\cdot}{\sim} N_K\{\sqrt{n}\underset{\sim}{\rho}, V\}.$$

Under H_o, $\underset{\sim}{\rho}=0$ and V=I, hence the null likelihood is given by

$$\begin{aligned} L_o &= f_o(\sqrt{n}\,r_1, \sqrt{n}\,r_2, \ldots, \sqrt{n}\,r_K) \\ &= \prod_{j=1}^{K} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}r_j^2} \\ &= (2\pi)^{-K/2} \exp\{-\frac{n}{2}\sum_{j=1}^{K} r_j^2\} \end{aligned} \qquad (5.3.3)$$

Under $H_o \cup H_a$, the likelihood is given by

$$\begin{aligned} L &= f(\sqrt{n}\,r_1, \sqrt{n}\,r_2, \ldots, \sqrt{n}\,r_K) \\ &= \frac{|V|^{-\frac{1}{2}}}{(2\pi)^{K/2}} \exp\{-\frac{n}{2}(\underset{\sim}{r}-\underset{\sim}{\rho})' V^{-1}(\underset{\sim}{r}-\underset{\sim}{\rho})\} . \end{aligned} \qquad (5.3.4)$$

The likelihood ratio test is based on the ratio

$$\begin{aligned} \lambda = \frac{L_o}{L} &= \frac{(2\pi)^{-K/2}\exp\{-\frac{n}{2}\sum_{j=1}^{K} r_j^2\}}{(2\pi)^{-K/2}|V|^{-\frac{1}{2}}\exp\{-\frac{n}{2}(r-\rho)'V^{-1}(r-\rho)\}} \\ &= |V|^{-\frac{1}{2}}\exp\{-\frac{n}{2}\sum_{j=1}^{K} r_j^2 + \frac{n}{2}(\underset{\sim}{r}-\underset{\sim}{\rho})'V^{-1}(\underset{\sim}{r}-\underset{\sim}{\rho})\} . \qquad \square \end{aligned}$$

In any practical case, of course, the alternative value of Ø will be unknown. Since $\underset{\sim}{\rho}$ and V in equation (5.3.2) are functions of Ø, maximum likelihood estimates of $\underset{\sim}{\rho}$ and V must be obtained and the ratio

$$\hat{\lambda} = \frac{L_o}{L} = |\hat{V}|^{-\frac{1}{2}}\exp\{-\frac{n}{2}\sum_{j=1}^{K} r_j^2 + \frac{n}{2}(\underset{\sim}{r}-\hat{\underset{\sim}{\rho}})'\hat{V}^{-1}(\underset{\sim}{r}-\hat{\underset{\sim}{\rho}})\}$$

computed. The problem reduces to finding the maximum likelihood estimator of Ø, say $\hat{Ø}$. However, due to the complexity of $\hat{\lambda}$, $\hat{Ø}$ may be difficult to obtain. This problem is avoided if we utilize a well-known property of of maximum likelihood estimators, namely that if $\hat{\emptyset}$ is the maximum likelihood estimate (MLE) of $\emptyset$, then $\hat{\emptyset} \xrightarrow{P} \emptyset$ as $n \to \infty$. Hannan [16] shows that for $1 \le j \le K$, $r_j \xrightarrow{P} \rho_j$ as $n \to \infty$ (Hannan's result is actually proved for the much stronger case of almost sure convergence.) Since $\rho_J = \emptyset$ for model (5.3.1), we have that $r_J \xrightarrow{P} \emptyset$ as $n \to \infty$. Hence, replacing the MLE $\hat{\emptyset}$ by the J^{th} sample autocorrelation r_J in the likelihood ratio $\hat{\lambda}$ will yield an equivalent test asymptotically.

In Sections 5.3.2 and 5.3.3, we consider likelihood ratio tests for these special cases of model (5.3.1): (J=5, K=2) and (J=K).

5.3.2 Special Case: J=5, K=2

We first wish to consider the form of the likelihood ratio test when the lag J of model (5.3.1) is larger than K. Recall that this is the case in which the researcher has underestimated the true lag of the alternative error model. We choose J=5 and K=2 in order to ensure independence of the sample autocorrelations. Since V of equation (5.3.2) is diagonal, computation of the likelihood is simplified considerably. Our result is given in Theorem 15.

Theorem 15: The likelihood ratio test of H_o: $\emptyset = 0$ versus H_a: $\emptyset \neq 0$ for J=5 and K=2 in model (5.3.1) is equivalent to the Box-Pierce test: reject H_o if $n \sum_{j=1}^{2} r_j^2$ is too large.

Proof: For the case J=5 and K=2, we apply Lemma 5 and 6 to obtain $\underset{\sim}{\rho} = (0,0)'$ and

$$V = \begin{bmatrix} \frac{1+\emptyset^2}{1-\emptyset^2} & 0 \\ 0 & \frac{1+\emptyset^2}{1-\emptyset^2} \end{bmatrix} .$$

Next, compute

$$V^{-1} = \begin{bmatrix} \frac{1-\emptyset^2}{1+\emptyset^2} & 0 \\ 0 & \frac{1-\emptyset^2}{1+\emptyset^2} \end{bmatrix}$$

Substituting these quantities into equations (5.3.3) and (5.3.4) of Lemma 18, we obtain

$$L_o = (2\pi)^{-1}\exp\{-\frac{n}{2}\,(r_1^2+r_2^2)\}$$

$$L = (2\pi)^{-1}(\frac{1-\emptyset^2}{1+\emptyset^2})\exp\{-\frac{n}{2}(\frac{1-\emptyset^2}{1+\emptyset^2})(r_1^2+r_2^2)\}\ .$$

Since the form of L is workable, we proceed with finding the MLE of $\emptyset$, or equivalently $\xi = \dfrac{1-\emptyset^2}{1+\emptyset^2}$. Now

$$\ell n\ L = -\ell n(2\pi) + \ell n(\xi) - \frac{n}{2}\,\xi(r_1^2+r_2^2)$$

and

$$\frac{d(\ell nL)}{d\xi} = \frac{1}{\xi} - \frac{n}{2}\,(r_1^2+r_2^2)\quad .$$

Setting the derivative equal to zero, and solving for $\hat{\xi}$ we obtain $\hat{\xi} = \dfrac{2}{n(r_1+r_2)}$. Substituting $\hat{\xi}$ for ξ in L, we obtain the likelihood ratio

$$\hat{\lambda} = \frac{L_o}{L} = \frac{(2\pi)^{-1}\exp\{-\frac{n}{2}\,(r_1^2+r_2^2)\}}{(2\pi)^{-1}(\hat{\xi})\exp\ -\frac{n}{2}(\hat{\xi})(r_1^2+r_2^2)\}}$$

$$= \frac{(\hat{\xi})^{-1}\exp\{-\frac{n}{2}(r_1^2+r_2^2)\}}{\exp\{-\frac{n}{2}[\frac{2}{n(r_1^2+r_2^2)}](r_1^2+r_2^2)\}}$$

$$= \frac{n(r_1^2+r_2^2)}{2}\ \exp\{1-\frac{n}{2}(r_1^2+r_2^2)\}$$

Since $\hat{\lambda} > C^*$ implies $n(r_1^2+r_2^2) > C$, the likelihood ratio test reduces the Box-Pierce test: reject H_o when $n(r_1^2+r_2^2) > C$, where C is determined to make the test level α. □

This result lends support to our simulation study of Chapter 4. Recall that the lone case in which the powers of the Box-Pierce test dominated the power of the max-χ^2 test was one for which K is less than J (K=4, J=5). Since the likelihood ratio test is one for which the power is maximized for fixed significance level α, the simulation results are substantiated.

Since we hope to choose K larger than the hypothesized lag J in the alternative error model, a more important case to investigate the form of the likelihood ratio is one for which $J<K$. However, λ becomes very unwieldy when J is much less than K, since the possibility of dependent sample autocorrelations increases. In the following section, we consider the special case $(J=K)$.

5.3.3 Special Case: J=K

Consider now the case in which the number of sample autocorrelations included in the test statistic (K) is equal to the lag (J) of the alternative error model. The following theorem gives the form of the likelihood ratio when J=K.

Theorem 16: The likelihood ratio statistic for testing H_o: $\emptyset = 0$ versus H_a: $\emptyset \neq 0$ when J=K in model (5.3.1) is of the form

$$-2\ell n\hat{\lambda} = -\ell n\; g(r_K) + nr_K^2 \left\{ 1 + \frac{\underset{\sim}{r}'_{(K-1)}[I-(V^*)^{-1}]\underset{\sim}{r}_{(K-1)}}{r_K^2} \right\}$$

where $g(r_K)$ is some function of r_K not dependent on n, V^* is a matrix with ijth element $(V^*)_{ij} = f_{ij}(r_K)$, each of the f_{ij} independent of n, and $\underset{\sim}{r}'_{(K-1)} = (r_1, r_2, r_3, \ldots, r_{K-1})$.

Proof: From Lemmas 5 and 6, we obtain for the case J=K

$$\underset{\sim}{\rho}_{(K)} = (0, 0, 0, \ldots, 0, \emptyset)',$$

and

$$V_{(K)} = \begin{bmatrix} V^*_{(K-1)} & \underset{\sim}{0}_{(K-1)} \\ \underset{\sim}{0}'_{(K-1)} & 1-\emptyset^2 \end{bmatrix}$$

where V is KxK, V^* is $(K-1)x(K-1)$, $\underset{\sim}{0}'_{(K-1)} = (0,0,0, \ldots, 0)$ and the ij^{th} element of V^* is $(V^*)_{ij} = f_{ij}(\emptyset)$.

For matrices of the form of V, it is easy to show that

$$V^{-1}_{(K)} = \begin{bmatrix} (V^*_{(K-1)})^{-1} & \underset{\sim}{0}_{(K-1)} \\ \underset{\sim}{0}'_{(K-1)} & \dfrac{1}{1-\phi^2} \end{bmatrix}$$

and $|V_{(K)}| = |V^*_{(K-1)}|\cdot(1-\phi^2)$. Applying Lemma 18, we obtain

$$L_o = (2\pi)^{-K/2}\exp\{-\frac{n}{2}\sum_{j=1}^{K} r_j^2\}.$$

$$L = (2\pi)^{-K/2}|V^*_{(K-1)}|^{-\frac{1}{2}}(1-\phi^2)^{-\frac{1}{2}}\exp\{-\frac{n}{2}[\underset{\sim}{r}'_{(K-1)}(V^*)^{-1}\underset{\sim}{r}_{(K-1)} + \frac{(r_K-\phi)^2}{1-\phi^2}]\}$$

where $\underset{\sim}{r}'_{(K-1)} = (r_1, r_2, r_3, \ldots, r_{K-1})$.

Since the maximum likelihood estimate of ϕ will be very difficult to obtain given the form of L above, we will use the Yule-Walker estimate $\hat{\phi} = r_J$. Since in this case J=K, we have $\hat{\phi} = r_K$. Again, we point out that r_K is a consistent estimator of ϕ, as is the maximum likelihood estimate. Thus, the likelihood ratio test using the Yule-Walker estimate will, for large n, be equivalent to the likelihood ratio test based on MLE. Substituting r_K for ϕ in L, we form the ratio

$$\hat{\lambda} = \frac{L_o}{\hat{L}} = \frac{(2\pi)^{-K/2}\exp\{-\frac{n}{2}\sum_{j=1}^{K} r_j^2\}}{(2\pi)^{-K/2}|\hat{V}^*|^{-\frac{1}{2}}(1-r_K^2)^{-\frac{1}{2}}\exp\{-\frac{n}{2}[\underset{\sim}{r}'_{(K-1)}(\hat{V}^*)^{-1}\underset{\sim}{r}_{(K-1)}]\}}$$

$$= |\hat{V}^*|^{\frac{1}{2}}(1-r_K^2)^{\frac{1}{2}}\exp\{-\frac{n}{2}\sum_{j=1}^{K} r_j^2 + \frac{n}{2}[\underset{\sim}{r}'_{(K-1)}(V^*)^{-1}\underset{\sim}{r}_{(K-1)}]\} ,$$

where

$$(\hat{V}^*)_{ij} = f_{ij}(r_K).$$

Taking logs of both sides and multiplying by -2, we obtain

$$
\begin{aligned}
-2\ell n\hat{\lambda} &= -\ell n|\hat{V}^*| - \ell n(1-r_K^2) + n\sum_{j=1}^{K} r_j^2 - n[\underset{\sim}{r}'_{(K-1)}(\hat{V}^*)^{-1}\underset{\sim}{r}_{(K-1)}] \\
&= -\ell n|\hat{V}^*| - \ell n(1-r_K^2) + nr_K^2 + n\underset{\sim}{r}'_{(K-1)}\underset{\sim}{r}_{(K-1)} - n[\underset{\sim}{r}'_{(K-1)}(\hat{V}^*)^{-1}\underset{\sim}{r}_{(K-1)}] \\
&= -\ell n\{|V^*|(1-r_K)\} + nr_K^2\left\{1 + \frac{\underset{\sim}{r}'_{(K-1)}\underset{\sim}{r}_{(K-1)} - \underset{\sim}{r}'_{(K-1)}(\hat{V}^*)^{-1}\underset{\sim}{r}_{(K-1)}}{r_K^2}\right\} \\
&= -\ell ng(r_K) + nr_K^2\left\{1 + \frac{\underset{\sim}{r}'_{(K-1)}[I-(\hat{V}^*)^{-1}]\underset{\sim}{r}_{(K-1)}}{r_K^2}\right\} \quad . \qquad \square
\end{aligned}
$$

Note that the dominant term in the likelihood ratio of equation (5.3.5) appears to be nr_K^2. Since, in general, $\underset{\sim}{r} \xrightarrow{P} \underset{\sim}{\rho}$ as $n \to \infty$, we have for J=K, $\underset{\sim}{r}_{(K-1)} \xrightarrow{P} \underset{\sim}{0}_{(K-1)}$ and $r_K \xrightarrow{P} \emptyset$ as $n \to \infty$. Thus, for large n,

$$nr_K^2\left\{1 + \frac{\underset{\sim}{r}'_{(K-1)}[I-(\hat{V}^*)^{-1}]\underset{\sim}{r}_{(K-1)}}{r_K^2}\right\} \sim nr_K^2 \quad .$$

Similarly, for the max-χ^2 statistic $T^{(m)}$, we have

$$T^{(m)} = n \max_{1\leq j\leq K} \{r_j^2\} \sim nr_K^2$$

for large n. Hence, the likelihood ratio test and max-χ^2 test are asymptotically equivalent. In order to more clearly see the dominance of the nr_K^2 term in the likelihood ratio, we consider a simple case: J=K=2. The result follows in the corollary.

Corollary 4: The likelihood ratio statistic for testing H_o: $\emptyset = 0$ versus H_a: $\emptyset \neq 0$ for J=K=2 in model (5.3.1) is given by,

$$-2\ell n\hat{\lambda} = -2\ell n(1+r_2) + nr_2^2\left\{1 + \frac{2r_1^2}{r_2(1+r_2)}\right\} \qquad (5.3.6).$$

Proof: We have previously determined $\underset{\sim}{\rho}$ and V for this case in Section 3.3.3.1. These expressions are given by,

$$\underset{\sim}{\rho} = (0,\phi)' \quad , V = \begin{bmatrix} \frac{(1+\phi)^2}{1-\phi^2} & 0 \\ 0 & 1-\phi \end{bmatrix}$$

and

$$V^{-1} = \begin{bmatrix} \frac{1-\phi}{(1+\phi)} & 0 \\ 0 & \frac{1}{1-\phi} \end{bmatrix} \quad .$$

Applying Lemma 18, we obtain

$$L_o = (2\pi)^{-1}\exp\{-\frac{n}{2}(r_1^2+r_2^2)\}$$

and

$$L = (2\pi)^{-1}(1+\phi)^{-1}\exp\{-\frac{n}{2}[r_1^2\frac{(1-\phi^2)}{(1+\phi^2)} + \frac{(r_2-\phi)^2}{1-\phi^2}]\} \quad .$$

Using the Yule-Walker estimate $\hat{\phi} = r_J$, or in the case at hand, $\hat{\phi} = r_2$, we substitute r_2 for ϕ in L and form the ratio

$$\hat{\lambda} = \frac{L_o}{\hat{L}} = \frac{(2\pi)^{-1}\exp\{-\frac{n}{2}(r_1^2+r_2^2)\}}{(2\pi)^{-1}(1+r_2)^{-1}\exp\{-\frac{n}{2}[r_1^2\frac{(1-r_2^2)}{(1+r_2)^2}+0]\}}$$

$$= (1+r_2)\exp\{-\frac{n}{2}(r_1^2+r_2^2) + \frac{n}{2}r_1^2[\frac{1-r_2^2}{(1+r_2)^2}]\}$$

Now

$$\ell n\hat{\lambda} = \ell n(1+r_2) - \frac{n}{2}(r_1^2+r_2^2) + \frac{n}{2}r_1^2\frac{(1-r_2^2)}{(1+r_2)^2}$$

and thus,

$$-2\ell n\hat{\lambda} = -2\ell n(1+r_2) + n(r_1^2+r_2^2) - nr_1^2(1-r_2^2)/(1+r_2)^2$$

$$= -2\ell n(1+r_2) + nr_2^2 \left\{ 1 + \frac{r_1^2}{(1+r_2)^2} + \frac{r_1^2}{r_2^2} - \frac{r_1^2}{r_2^2(1+r_2)^2} \right\}$$

$$= -2\ell n(1+r_2) + nr_2^2 \left\{ 1 + \frac{r_1^2r_2^2 + r_1^2(1+r_2)^2 - r_1^2}{r_2(1+r_2)^2} \right\}$$

$$= -2\ell n(1+r_2) + nr_2^2 \left\{ 1 + \frac{2r_1^2r_2^2 + 2r_1^2r_2^2}{r_2^2(1+r_2)^2} \right\}$$

$$= -2\ell n(1+r_2) + nr_2^2 \left\{ 1 + \frac{2r_1^2r_2(r_2+1)}{r_2^2(1+r_2)^2} \right\}$$

$$= -2\ell n(1+r_2) + nr_2^2 \left\{ 1 + \frac{2r_1^2}{r_2(1+r_2)} \right\} \qquad \square$$

The form of the likelihood ratio in (5.3.6) is more easily seen to be dominated by nr_2^2 for large n. Again, when K=J=2, $T^{(m)} = n \max(r_1^2, r_2^2) \sim nr_2$ for large n, and the test statistics are in agreement asymptotically.

One need only examine equation (5.3.6) for the simple case (J=2,K=2) to grasp the awkwardness of the likelihood ratio statistic. For those cases where K is much larger than J, the form becomes even more unwieldy than equation (5.3.5). Although the fact that $-2\ell n\hat{\lambda} \xrightarrow[H_o]{D} \chi_1^2$ will enable the researcher to obtain asymptotic critical regions for the likelihood ratio, the exact form of the statistic will change for each different (J,K). For practical research work then, the likelihood ratio test is unfeasible.

The form of the ratio does, however, provide an insight into the power simulation results for the Box-Pierce and max-χ^2 tests. For a particular case of K<J, the likelihood ratio is equivalent to the Box-Pierce test, while for K=J, the dominant term of the ratio is the max-χ^2 statistic. Although the case J<K yields an intractible likelihood ratio,

our power simulation results hint that some form of the max-χ^2 statistic will be the dominant term in the likelihood ratio.

In conclusion, we state the main results of this chapter.

(i) Asymptotic power comparisons indicate that for the first order AR lag J model, the max-χ^2 test attains a power of 1 at a faster rate than the Box-Pierce test.

(ii) The HLARE of the Box-Pierce test to the max-χ^2 test is 1, i.e., the limiting ratio of the sample sizes needed for the tests to have equivalent power is 1.

(iii) The Box-Pierce test was found to be asymptotically deficient, in the sense that the number of additional observations needed for the power of the B-P test to be equivalent to the max-χ^2 power tends to infinity as n grows large.

(iv) The likelihood ratio for J=K is dominated by a form of the max-χ^2 statistic, providing further support for the dominance of the max-χ^2 test in the power simulation results of Chapter 4.

CHAPTER VI

CONCLUSION

6.1 Concluding Remarks

The aim of this research has been to present a new test designed to detect serial correlation of the errors in least squares regression models. The problem of testing for dependent regression errors has been widely researched since the early 1950's; however, most of the procedures developed attain high power only when the nature of the residual correlation is first order autoregressive. Our new procedure--the max-χ^2 technique--is designed to attain high power against a general p^{th} order autoregressive error model.

A substantial portion of this research has been devoted to a comparison of the max-χ^2 test with an alternative procedure also designed to detect serial correlation in the general case--the Box-Pierce test. An extensive Monte-Carlo simulation study was undertaken to compare the powers of the two tests when the regression errors under the alternative hypothesis of serial dependence follow the first order lag J autoregressive process:

$$Z_t = \emptyset Z_{t-J} + \varepsilon_t \quad ,$$

where $\{\varepsilon_t\}$ is white noise. Our reason for examining this model is that it represents one which is close to the null hypothesis of white noise, and thus one for which many testing procedures will fail to detect serial correlation. We hypothesized that under this model (1) both the max- χ^2

and Box-Pierce tests would have high power and (2) that the max-χ^2 procedure would outperform the Box-Pierce test. Our simulation results appear to strongly support this conjecture. In addition to the Box-Pierce and max-χ^2 tests, we also gave consideration to the well-known and widely used Durbin-Watson d test. The Durbin-Watson test performed extremely well when J=1 in the AR model, a result supported by other researchers. However, for J>1, the d test lacked the high power of the two competing tests.

In practice, our recommendations to the researcher who seeks to detect serial correlation in low order AR models are:

(i) Apply the Durbin-Watson d test initially, since it has high power against the AR lag 1 model, a recurrent and hence most important case in practice.

(ii) If the d test either fails to reject the null hypothesis of independent errors or falls into its "inconclusive region", apply the max-χ^2 procedure, a test with high power against a first order AR error scheme with lag greater than 1.

We note that the researcher should be concerned with the overall significance level of the testing procedure, since the two tests are not independent. We suggest the Bonferroni method of performing each test at significance level $\alpha/2$, for this guarantees a true overall significance level which is less than or equal to α.

6.2 Future Research

Possible avenues for research in the area of testing for serial correlation are limitless. Below, we give three main areas which arise naturally from the work presented in this dissertation.

(1) Although our Monte-Carlo simulation study was extensive, it was by no means exhaustive. We basically considered a first order autoregressive error model with a single non-zero lag parameter (the lag J not necessarily 1), for this model is close to the null hypothesis of white noise and most vulnerable to being undetected. However, it would be interesting to note what effect a general autoregressive-moving average time series alternative error model would have on the powers of the tests. An extension of our Monte-Carlo simulation to the general ARMA case is one possible avenue of future research.

(2) An important discovery of the simulation study was the delicateness of the researcher's choice of K, the number of sample autocorrelations to be included in the test statistic. A value of K much larger than the maximum lag in the alternative error model will effectively reduce the power of the test, while K less than the minimum lag essentially makes the test powerless. McClave's max-χ^2 method was originally designed to estimate the order (J) of autoregressive models, and it may be worthwhile for the researcher to apply this method first, and base his choice of K on the estimated order. Another possibility, from a Bayesian view, would be to attache a prior probability on the possible values of J. Since high power results when K is nearly equal to J, the mean of the posterior distribution of J may be of value in determining a suitable K. A method for choosing K, then, is a very important topic in future research of the problem.

(3) We have stated that our inability to compute asymptotic powers of the Box-Pierce test for large K was due, in part, to the near singularities of the sample autocorrelation covariance matrix (V). Most of the eigenvalues of V in this instance are very small (near zero) and the

remaining few considerably larger. This phenomenon of a near singular covariance matrix may very well be the reason why the power of the Box-Pierce test diminishes as K is increased. This problem was not as much a factor in the max-χ^2 case, since only the largest sample autocorrelation is considered. Hence it may be worthwhile to examine tests based on the eigenvalues of the sample covariance matrix of the sample autocorrelations. One possibility would be to use as a test statistic the maximum eigenvalue of the estimated covariance matrix (V) of the sample autocorrelations. An alternate test to consider is that which is based on the linear combination of the sample autocorrelations associated with the largest eigenvalue of V. From principal components analysis, we know that the coefficients of the linear combination are simply the elements of the eigenvector associated with the largest eigenvalue. Since asymptotic properties of principal components are well known, the null distribution of the test statistic is tractible for large n.

APPENDIX

TABLE A-1

POWER SIMULATIONS: $J=2,\ K=4,\ \alpha=.05$
ERROR MODEL 1: $Z_t=\phi Z_{t-2}+\varepsilon_t$

	n=50		n=200	
ϕ	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-.9	1.000	1.000	1.000	1.000
-.7	.987	.992	1.000	1.000
-.5	.764	.817	1.000	1.000
-.3	.290	.322	.937	.960
-.1	.048	.050	.125	.139
0	.050	.045	.049	.052
.1	.068	.055	.160	.145
.3	.333	.316	.947	.959
.5	.798	.814	1.000	1.000
.7	.989	.993	1.000	1.000
.9	1.000	1.000	1.000	1.000

TABLE A-2

POWER SIMULATIONS: $J=3$, $K=4$, $\alpha=.05$
ERROR MODEL 1: $Z_t=\emptyset Z_{t-3}+\varepsilon_t$

	n=50		n=200	
Ø	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-.9	1.000	1.000	1.000	1.000
-.7	.976	.985	1.000	1.000
-.5	.755	.796	1.000	1.000
-.3	.304	.322	.937	.953
-.1	.064	.049	.157	.166
0	.038	.040	.049	.052
.1	.049	.047	.143	.147
.3	.258	.281	.948	.961
.5	.782	.826	1.000	1.000
.7	.976	.988	1.000	1.000
.9	1.000	1.000	1.000	1.000

TABLE A-3

POWER SIMULATIONS: J=4, K=4, α=.05
ERROR MODEL 1: $Z_t=\emptyset Z_{t-4}+\varepsilon_t$

	n=50		n=200	
$\emptyset$	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-.9	.999	.999	1.000	1.000
-.7	.971	.983	1.000	1.000
-.5	.741	.815	1.000	1.000
-.3	.229	.295	.926	.949
-.1	.052	.037	.118	.140
0	.038	.040	.056	.059
.1	.065	.044	.164	.156
.3	.310	.303	.938	.958
.5	.760	.775	1.000	1.000
.7	.981	.984	1.000	1.000
.9	1.000	1.000	1.000	1.000

TABLE A-4

POWER SIMULATIONS: J=2, K=12, α=.05
ERROR MODEL 1: $Z_t = \phi Z_{t-2} + \varepsilon_t$

	n=50		n=200	
Ø	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-.9	.999	.999	1.000	1.000
-.7	.937	.982	1.000	1.000
-.5	.555	.673	1.000	1.000
-.3	.175	.184	.793	.916
-.1	.024	.018	.087	.080
0	.034	.022	.043	.037
.1	.035	.018	.115	.096
.3	.199	.173	.837	.925
.5	.616	.718	1.000	1.000
.7	.948	.969	1.000	1.000
.9	1.000	1.000	1.000	1.000

TABLE A-5

POWER SIMULATIONS: $J=3$, $K=12$, $\alpha=.05$
ERROR MODEL 1: $Z_t=\phi Z_{t-3}+\varepsilon_t$

	n=50		n=200	
ϕ	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-.9	.998	1.000	1.000	1.000
-.7	.930	.975	1.000	1.000
-.5	.577	.644	.999	1.000
-.3	.173	.156	.809	.907
-.1	.020	.014	.103	.097
0	.034	.022	.043	.037
.1	.034	.019	.106	.096
.3	.180	.173	.821	.919
.5	.575	.671	1.000	1.000
.7	.938	.963	1.000	1.000
.9	1.000	1.000	1.000	1.000

TABLE A-6

POWER SIMULATIONS: J=4, K=12, $\alpha=.05$
ERROR MODEL 1: $Z_t=\phi Z_{t-4}+\varepsilon_t$

	n=50		n=200	
ϕ	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-.9	.997	.999	1.000	1.000
-.7	.912	.963	1.000	1.000
-.5	.546	.664	1.000	1.000
-.3	.143	.154	.803	.909
-.1	.040	.012	.095	.076
0	.034	.022	.043	.037
.1	.040	.017	.134	.101
.3	.176	.168	.827	.914
.5	.583	.646	.999	1.000
.7	.937	.966	1.000	1.000
.9	.999	1.000	1.000	1.000

TABLE A-7

POWER SIMULATIONS: $J=12$, $K=12$, $\alpha=.05$
ERROR MODEL 1: $Z_t=\phi Z_{t-12}+\varepsilon_t$

	n=50		n=200	
ϕ	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-.9	.992	1.000	1.000	1.000
-.7	.755	.910	1.000	1.000
-.5	.307	.394	1.000	1.000
-.3	.096	.053	.736	.860
-.1	.028	.009	.086	.089
0	.034	.022	.043	.037
.1	.036	.012	.122	.098
.3	.122	.079	.787	.884
.5	.394	.406	.998	1.000
.7	.823	.895	1.000	1.000
.9	.996	1.000	1.000	1.000

TABLE A-8

POWER SIMULATIONS: $J=12$, $K=24$, $\alpha=.05$
ERROR MODEL 1: $Z_t=\emptyset Z_{t-12}+\varepsilon_t$

	n=50		n=200	
$\emptyset$	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-.9	.976	1.000	1.000	1.000
-.7	.583	.853	1.000	1.000
-.5	.196	.278	.997	1.000
-.3	.048	.024	.614	.808
-.1	.012	.008	.063	.058
0	.014	.004	.041	.027
.1	.009	.004	.083	.071
.3	.071	.039	.648	.843
.5	.259	.298	.994	1.000
.7	.664	.837	1.000	1.000
.9	.977	1.000	1.000	1.000

TABLE A-9

POWER SIMULATIONS: J=5, K=4, α=.05
ERROR MODEL 1: $Z_t=\phi Z_{t-5}+\varepsilon_t$

n=50

ϕ	Box-Pierce	Max-χ^2
-.9	.466	.339
-.7	.292	.192
-.5	.135	.085
-.3	.075	.051
-.1	.041	.033
0	.058	.048
.1	.047	.035
.3	.082	.057
.5	.157	.104
.7	.287	.194
.9	.454	.331

TABLE A-10

POWER SIMULATIONS: $J_1=1,\ J_2=2,\ K=2,\ \alpha=.05$
ERROR MODEL 2: $Z_t=\emptyset_t Z_{t-1}+\emptyset_2 Z_{t-2}+\varepsilon_t$

		n=50		n=200	
$\emptyset_1$	$\emptyset_2$	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-1.0	-.6	1.000	1.000	1.000	1.000
- .7	.3	1.000	1.000	1.000	1.000
- .5	.3	.993	.986	1.000	1.000
- .3	.1	.584	.568	.989	.987
0	0	.042	.042	.056	.053
.1	-.7	.995	.997	1.000	1.000
.1	.7	.995	.995	1.000	1.000
.5	.3	.992	.989	1.000	1.000
.7	.3	1.000	1.000	1.000	1.000
1.0	-.3	1.000	1.000	1.000	1.000
1.0	-.6	1.000	1.000	1.000	1.000

TABLE A-11

POWER SIMULATIONS: $J_1=1$, $J_2=2$, $K=12$, $\alpha=.05$
ERROR MODEL 2: $Z_t=\emptyset_1 Z_{t-1}+\emptyset_2 Z_{t-2}+\varepsilon_t$

		n=50		n=200	
$\emptyset_1$	0_2	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-1.0	-.6	.987	.996	1.000	1.000
- .7	.3	1.000	1.000	1.000	1.000
- .5	.3	.960	.950	1.000	1.000
- .3	.1	.323	.307	.925	.967
0	0	.034	.022	.043	.037
.1	-.7	.943	.975	1.000	1.000
.1	.7	.966	.992	1.000	1.000
.5	.3	.953	.951	1.000	1.000
.7	.3	1.000	1.000	1.000	1.000
1.0	-.3	.996	1.000	1.000	1.000
1.0	-.6	.987	.996	1.000	1.000

TABLE A-12

POWER SIMULATIONS: $J_1=1,\ J_2=3,\ K=4,\ \alpha=.05$
ERROR MODEL 2: $Z_t=\phi_1 Z_{t-1}+\phi_3 Z_{t-3}+\varepsilon_t$

		n=50		n=200	
ϕ_1	ϕ_2	Box-Pierce	Max-χ^2	Box-Pierce	Max-χ^2
-1.0	.3	1.000	1.000	1.000	1.000
-1.0	.1	.999	1.000	1.000	1.000
- .5	- .3	.961	.942	1.000	1.000
- .1	.3	.301	.300	.969	.971
0	0	.038	.040	.049	.052
.1	- .9	1.000	1.000	1.000	1.000
.1	.7	.983	.992	1.000	1.000
.3	- .1	.346	.308	.956	.965
.5	- .3	.903	.880	1.000	1.000
.5	-1.2	1.000	1.000	1.000	1.000
1.0	.1	1.000	1.000	1.000	1.000

TABLE A-13

POWER SIMULATIONS: $J=2$, $K=4$, $\alpha=.05$
REGRESSION MODEL: $Y_t=.4+4.4t+Z_t$
ERROR MODEL: $Z_t=\emptyset Z_{t-2}+\varepsilon_t$

n=50

Ø	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	.999	1.000	.000	.000
-.8	1.000	1.000	.000	.001
-.7	.987	.996	.000	.000
-.6	.929	.958	.000	.001
-.5	.808	.875	.001	.000
-.4	.564	.637	.001	.006
-.3	.338	.394	.006	.010
-.2	.152	.160	.008	.018
-.1	.061	.067	.026	.018
0	.028	.033	.036	.035
.1	.054	.050	.059	.054
.2	.123	.096	.081	.057
.3	.287	.235	.128	.072
.4	.499	.475	.185	.088
.5	.691	.704	.243	.072
.6	.860	.852	.307	.081
.7	.955	.958	.409	.085
.8	.990	.990	.549	.053
.9	.995	.995	.703	.037

TABLE A-14

POWER SIMULATIONS: $J=3$, $K=4$, $\alpha=.05$
REGRESSION MODEL: $Y_t=.4+4.4t+Z_t$
ERROR MODEL: $Z_t=\phi Z_{t-3}+\varepsilon_t$

$n=50$

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	.346	.082
-.8	.998	.999	.222	.081
-.7	.990	.997	.140	.072
-.6	.926	.960	.114	.057
-.5	.797	.851	.093	.051
-.4	.586	.640	.059	.053
-.3	.309	.317	.037	.038
-.2	.190	.182	.046	.032
-.1	.091	.081	.044	.024
0	.028	.033	.036	.035
.1	.067	.051	.052	.054
.2	.119	.095	.034	.037
.3	.244	.222	.042	.049
.4	.446	.415	.054	.053
.5	.644	.655	.077	.072
.6	.834	.851	.116	.096
.7	.952	.951	.179	.117
.8	.989	.988	.356	.137
.9	.999	.999	.584	.108

TABLE A-15

POWER SIMULATIONS: $J=4$, $K=4$, $\alpha=.05$
REGRESSION MODEL: $Y_t=.4+4.4t+Z_t$
ERROR MODEL: $Z_t=\phi Z_{t-4}+\varepsilon_t$

$n=50$

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	.351	.071
-.8	.997	1.000	.236	.089
-.7	.984	.992	.174	.087
-.6	.926	.959	.122	.064
-.5	.774	.840	.080	.053
-.4	.520	.591	.060	.062
-.3	.306	.363	.062	.034
-.2	.130	.137	.045	.047
-.1	.059	.053	.043	.028
0	.028	.033	.036	.035
.1	.071	.048	.037	.032
.2	.149	.103	.050	.048
.3	.273	.218	.056	.033
.4	.459	.422	.062	.048
.5	.683	.654	.080	.050
.6	.866	.871	.124	.061
.7	.945	.945	.186	.057
.8	.993	.992	.247	.066
.9	1.000	1.000	.338	.058

TABLE A-16

POWER SIMULATIONS: J=1, K=12, α=.05
REGRESSION MODEL: $Y_t = .4+4.4t+Z_t$
ERROR MODEL: $Z_t = \phi Z_{t-1} + \varepsilon_t$

n=50

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	.999	1.000	1.000	.000
-.8	.994	.999	1.000	.000
-.7	.959	.990	.995	.003
-.6	.869	.934	.974	.007
-.5	.664	.764	.909	.036
-.4	.452	.528	.779	.081
-.3	.230	.263	.505	.127
-.2	.123	.105	.286	.089
-.1	.055	.044	.095	.061
0	.028	.019	.030	.070
.1	.033	.020	.048	.042
.2	.077	.045	.188	.086
.3	.165	.130	.408	.106
.4	.331	.313	.661	.094
.5	.508	.545	.862	.055
.6	.724	.796	.953	.020
.7	.869	.920	.988	.006
.8	.947	.975	.996	.004
.9	.983	.992	1.000	.000

TABLE A-17

POWER SIMULATIONS: $J=4$, $K=12$, $\alpha=.05$
REGRESSION MODEL: $Y_t=.4+4.4t+Z_t$
ERROR MODEL: $Z_t=\emptyset Z_{t-4}+\varepsilon_t$

n=50

∅	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	.998	.999	.338	.076
-.8	.977	.992	.252	.081
-.7	.928	.973	.188	.057
-.6	.765	.905	.113	.069
-.5	.586	.721	.084	.063
-.4	.359	.466	.074	.057
-.3	.192	.215	.042	.040
-.2	.090	.073	.049	.042
-.1	.047	.028	.042	.037
0	.028	.019	.030	.070
.1	.054	.031	.036	.036
.2	.102	.061	.064	.031
.3	.180	.129	.046	.044
.4	.331	.292	.072	.059
.5	.548	.528	.097	.051
.6	.742	.753	.144	.055
.7	.908	.926	.181	.057
.8	.968	.973	.234	.073
.9	.989	.996	.332	.058

TABLE A-18

POWER SIMULATIONS: $J=4$, $K=4$, $\alpha=.05$
REGRESSION MODEL: $Y_t=21.548-.2544t+.0527t^2+Z_t$
ERROR MODEL: $Z_t=\emptyset Z_{t-4}+\varepsilon_t$

n=50

$\emptyset$	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	.999	.999	.332	.147
-.8	.999	1.000	.194	.156
-.7	.982	.991	.150	.148
-.6	.925	.960	.083	.143
-.5	.790	.839	.064	.118
-.4	.546	.642	.049	.093
-.3	.330	.370	.050	.079
-.2	.152	.172	.022	.083
-.1	.078	.080	.033	.071
0	.067	.058	.032	.043
.1	.081	.080	.029	.085
.2	.149	.106	.037	.107
.3	.295	.221	.061	.095
.4	.509	.431	.070	.098
.5	.717	.637	.099	.112
.6	.865	.835	.133	.153
.7	.963	.954	.205	.131
.8	.991	.992	.283	.135
.9	.999	.999	.358	.118

TABLE A-19

POWER SIMULATIONS: $J=1$, $K=12$, $\alpha=.05$
REGRESSION MODEL: $Y_t=21.548-.2544t+.05727t^2+Z_t$
ERROR MODEL: $Z_t=\phi Z_{t-1}+\varepsilon_t$
$n=50$

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	.999	1.000	1.000	.000
-.8	.995	.999	.999	.001
-.7	.955	.990	.995	.003
-.6	.881	.944	.974	.020
-.5	.667	.814	.992	.057
-.4	.438	.552	.751	.144
-.3	.247	.307	.522	.224
-.2	.130	.114	.240	.222
-.1	.065	.056	.077	.144
0	.047	.030	.024	.073
.1	.047	.025	.031	.082
.2	.091	.056	.105	.158
.3	.180	.115	.271	.201
.4	.333	.260	.505	.233
.5	.535	.489	.768	.121
.6	.699	.690	.860	.085
.7	.841	.844	.962	.022
.8	.938	.946	.984	.011
.9	.961	.964	.994	.003

TABLE A-20

POWER SIMULATIONS: J=4, K=12, α=.05
REGRESSION MODEL: $Y_t=21.548-.2544t+.05727t^2+Z_t$
ERROR MODEL: $Z_t=\phi Z_{t-4}+\varepsilon_t$

n=50

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	.320	.158
-.8	.985	.997	.201	.155
-.7	.922	.973	.149	.149
-.6	.792	.885	.122	.139
-.5	.573	.720	.060	.111
-.4	.374	.459	.042	.128
-.3	.189	.215	.037	.090
-.2	.110	.093	.029	.098
-.1	.059	.040	.039	.069
0	.047	.030	.024	.073
.1	.064	.038	.035	.081
.2	.112	.054	.030	.081
.3	.209	.133	.042	.081
.4	.373	.279	.069	.128
.5	.608	.516	.101	.122
.6	.781	.748	.150	.152
.7	.923	.908	.186	.135
.8	.973	.974	.251	.143
.9	.997	.998	.333	.125

TABLE A-21

POWER SIMULATIONS: J=2, K=4, $\alpha=.05$

REGRESSION MODEL: $Y_t=.84+.098X_t-.00057t+Z_t$

$X_t=.8+.8X_{t-1}+\eta_t$

ERROR MODEL: $Z_t=\phi Z_{t-2}+\varepsilon_t$

n=50

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	.000	.000
-.7	.987	.996	.000	.001
-.5	.831	.876	.000	.006
-.3	.341	.407	.002	.029
-.1	.072	.074	.014	.070
0	.044	.042	.026	.084
.1	.061	.040	.042	.116
.3	.255	.214	.125	.130
.5	.665	.632	.217	.175
.7	.945	.946	.425	.127
.9	.992	.999	.747	.067

TABLE A-22

POWER SIMULATIONS: $J=4$, $K=4$, $\alpha=.05$
REGRESSION MODEL: $Y_t=.84+.098X_t-.00057t+Z_t$
$X_t=.8+.8X_{t-1}+\eta_t$
ERROR MODEL: $Z_t=\phi Z_{t-4}+\varepsilon_t$

$n=50$

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	.324	.169
-.7	.980	.987	.145	.174
-.5	.756	.824	.079	.123
-.3	.299	.334	.040	.100
-.1	.063	.059	.032	.082
0	.044	.042	.026	.084
.1	.062	.058	.031	.079
.3	.289	.233	.038	.089
.5	.700	.665	.077	.113
.7	.953	.954	.155	.117
.9	1.000	.999	.303	.109

TABLE A-23

POWER SIMULATIONS: $J=1$, $K=12$, $\alpha=.05$

REGRESSION MODEL: $Y_t=.84+.098X_t-.00057t+Z_t$

$X_t=.8+.8X_{t-1}+\eta_t$

ERROR MODEL: $Z_t=\phi Z_{t-1}+\varepsilon_t$

$n=50$

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	.999	.999	1.000	.000
-.8	.991	.998	.999	.000
-.7	.960	.988	.998	.000
-.6	.860	.935	.980	.016
-.5	.653	.783	.901	.061
-.4	.433	.527	.739	.141
-.3	.251	.286	.481	.229
-.2	.138	.105	.234	.212
-.1	.055	.036	.088	.138
0	.047	.032	.029	.083
.1	.037	.018	.044	.086
.2	.067	.039	.110	.182
.3	.133	.097	.300	.224
.4	.270	.244	.544	.188
.5	.460	.492	.752	.147
.6	.668	.720	.912	.054
.7	.845	.889	.968	.022
.8	.928	.959	.990	.007
.9	.965	.984	.997	.001

TABLE A-24

POWER SIMULATIONS: $J=4$, $K=12$, $\alpha=.05$
REGRESSION MODELS: $Y=.84+.098X-.00057t+Z_t$
$X_t=.8+.8X_{t-4}+\eta_t$

ERROR MODEL: $Z_t=\phi Z_{t-4}+\varepsilon_t$

$n=50$

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	.342	.146
-.8	.989	.998	.199	.183
-.7	.932	.977	.122	.175
-.6	.807	.902	.078	.160
-.5	.576	.729	.059	.119
-.4	.355	.457	.045	.098
-.3	.199	.202	.032	.101
-.2	.099	.074	.027	.091
-.1	.049	.039	.209	.083
0	.047	.032	.029	.083
.1	.063	.034	.029	.084
.2	.104	.054	.041	.079
.3	.204	.121	.049	.090
.4	.358	.277	.058	.102
.5	.533	.522	.078	.110
.6	.755	.748	.104	.106
.7	.912	.923	.147	.115
.8	.973	.979	.194	.131
.9	.996	.996	.282	.105

TABLE A-25

POWER SIMULATIONS: $J=2$, $K=4$, $\alpha=.05$

REGRESSION MODEL: $Y_t=2.14+.69X_{1t}-.63X_{2t}+Z_t$

$$\left.\begin{array}{l} X_{1t}=.8+.8X_{1,t-1}+\omega_{1t} \\ X_{2t}=.3-.5X_{2,t-1}+\omega_{2t} \end{array}\right\} \mathrm{Cov}(\omega_{1t},\omega_{2t})=.7$$

ERROR MODEL: $Z_t=\phi Z_{t-2}+\varepsilon_t$

$n=50$

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	.000	.000
-.8	.996	.997	.000	.000
-.7	.989	.992	.000	.000
-.6	.943	.962	.000	.002
-.5	.793	.857	.000	.006
-.4	.561	.634	.000	.015
-.3	.333	.373	.002	.025
-.2	.158	.171	.012	.026
-.1	.059	.068	.014	.048
0	.037	.035	.032	.034
.1	.062	.042	.038	.093
.2	.124	.093	.058	.111
.3	.250	.216	.094	.123
.4	.461	.437	.126	.143
.5	.701	.699	.185	.143
.6	.871	.874	.250	.142
.7	.959	.965	.329	.158
.8	.992	.994	.453	.125
.9	.998	.998	.577	.096

TABLE A-26

POWER SIMULATIONS: $J=4,\ K=4,\ \alpha=.05$

REGRESSION MODEL: $Y_t=2.14+.69X_{1t}-.63X_{1t}+Z_t$

$$\left.\begin{array}{l} X_{1t}=.8+.8X_{1,t-1}+\omega_{1t} \\ X_{2t}=.3-.5X_{2,t-1}+\omega_{2t} \end{array}\right\} \mathrm{Cov}(\omega_{1t},\omega_{2t})=.7$$

ERROR MODEL: $Z_t=\phi Z_{t-4}+\varepsilon_t$

$n=50$

ϕ	Box-Pierce	Max- χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	.348	.143
-.7	.971	.986	.148	.154
-.5	.726	.767	.065	.111
-.3	.245	.285	.035	.091
-.1	.055	.051	.208	.080
0	.037	.035	.032	.034
.1	.062	.040	.022	.085
.3	.274	.224	.025	.102
.5	.689	.672	.046	.119
.7	.954	.955	.100	.109
.9	.999	.999	.211	.096

TABLE A-27

POWER SIMULATIONS: $J=1$, $K=12$, $\alpha=.05$
REGRESSION MODEL: $Y_t=2.14+.69X_{1t}-.63X_{2t}+Z_t$

$$\left.\begin{array}{l} X_{1t}=.8+.8X_{1,t-1}+\omega_{1t} \\ X_{2t}=.3-.5X_{2,t-1}+\omega_{2t} \end{array}\right\} \mathrm{Cov}(\omega_{1t},\omega_{2t})=.7$$

ERROR MODEL: $Z_t=\phi Z_{t-1}+Z_t$

$n=50$

ϕ	Box-Pierce	Max-χ^2	Durbin-Watson	D-W Inconclusive
-.9	1.000	1.000	1.000	.000
-.8	.986	.993	1.000	.000
-.7	.942	.977	.985	.010
-.6	.839	.909	.959	.023
-.5	.633	.747	.869	.084
-.4	.399	.483	.690	.157
-.3	.216	.226	.436	.215
-.2	.104	.073	.182	.219
-.1	.051	.028	.054	.122
0	.030	.014	.017	.080
.1	.041	.018	.056	.090
.2	.071	.040	.150	.175
.3	.138	.116	.358	.231
.4	.273	.290	.624	.180
.5	.469	.570	.826	.106
.6	.690	.790	.948	.034
.7	.869	.926	.988	.011
.8	.957	.989	1.000	.000
.9	.991	.998	1.000	.000

TABLE A-28

POWER SIMULATIONS: $J=4,\ K=12,\ \alpha=.05$

REGRESSION MODEL: $Y_t=2.14+.69X_{1t}-.63X_{2t}+Z_t$

$$\left.\begin{array}{l} X_{1t}=.8+.8X_{1,t-1}+\omega_{1t} \\ X_{2t}=.3-.5X_{2,t-1}+\omega_{2t} \end{array}\right\} \mathrm{Cov}(\omega_{1t},\omega_{2t})=.7$$

ERROR MODEL: $Z_t=ØZ_{t-4}+\varepsilon_t$

n=50

Ø	Box-Pierce	Max-χ^2	Durbin-Watson	D-w Inconclusive
-.9	.999	1.000	.309	.145
-.8	.983	.996	.190	.160
-.7	.917	.968	.123	.145
-.6	.774	.864	.088	.135
-.5	.539	.665	.065	.125
-.4	.327	.415	.045	.105
-.3	.166	.179	.033	.088
-.2	.087	.069	.028	.079
-.1	.042	.022	.022	.074
0	.030	.014	.017	.080
.1	.035	.020	.020	.074
.2	.063	.038	.025	.076
.3	.150	.117	.209	.078
.4	.286	.293	.036	.093
.5	.505	.545	.048	.095
.6	.712	.776	.072	.096
.7	.877	.933	.101	.118
.8	.969	.986	.143	.120
.9	.996	.997	.220	.110

TABLE A-29

POWER APPROXIMATIONS: Box-Pierce Test
ERROR MODEL 1: J=2, K=4, α=.05

BP(1) = Imhoff numerical integration
BP(2) = Pearson 3-moment extension
BP(3) = Rao square root transformation
BP(4) = Simulated Power

	n=50				n=200			
Ø	BP(1)	BP(2)	BP(3)	BP(4)	BP(1)	BP(2)	BP(3)	BP(4)
-.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-.8	1.000	1.000	.999	1.000				
-.7	.998	1.000	.993	.987	1.000	1.000	1.000	1.000
-.6	.996	1.000	.963	.917				
-.5	.839	1.000	.765	.775	1.000	1.000	.999	1.000
-.4	.606	.874	.488	.552				
-.3	.352	.304	.218	.283	.944	1.000	.888	.937
-.2	.164	.097	.060	.114				
-.1	.070	.052	.010	.036	.163	.116	.051	.125
0	.050	.050	.002	.050	.050	.050	.002	.049
.1	.094	.064	.010	.062	.192	.140	.051	.160
.2	.225	.147	.060	.162				
.3	.451	.501	.218	.334	.958	1.000	.888	.947
.4	.708	1.000	.488	.588				
.5	.900	1.000	.765	.828	1.000	1.000	.999	1.000
.6	.983	1.000	.963	.946				
.7	.999	1.000	.993	.979	1.000	1.000	1.000	1.000
.8	1.000	1.000	.999	1.000				
.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE A-30

POWER APPROXIMATIONS: Max-χ^2 Test
ERROR MODEL 1: J=2, K=4, α=.05

M(1) = Taylor-series approximation M(2) = Simulated power

	n=50		n=200	
Ø	M(1)	M(2)	M(1)	M(2)
-.9	1.000	1.000	1.000	1.000
-.8	1.000	1.000		
-.7	.999	.994	1.000	1.000
-.6	.987	.948		
-.5	.889	.827	1.000	1.000
-.4	.653	.602		
-.3	.370	.320	.967	.960
-.2	.163	.118		
-.1	.068	.040	.167	.139
0	.050	.045	.050	.052
.1	.086	.057	.183	.145
.2	.198	.144		
.3	.415	.325	.970	.959
.4	.692	.573		
.5	.908	.814	1.000	1.000
.6	.990	.958		
.7	.999	.988	1.000	1.000
.8	1.000	1.000		
.9	1.000	1.000	1.000	1.000

TABLE A-31

POWER APPROXIMATIONS: Box-Pierce Test
ERROR MODEL 1: J=3, K=4, α=.05

BP(1) = Imhoff numerical integration
BP(2) = Pearson 3-moment extension
BP(3) = Rao square root transformation
BP(4) = Simulated power

	n=50				n=200			
ϕ	BP(1)	BP(2)	BP(3)	BP(4)	BP(1)	BP(2)	BP(3)	BP(4)
-.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-.8	1.000	1.000	.999	.998				
-.7	.999	1.000	.993	.976	1.000	1.000	1.000	1.000
-.6	.978	1.000	.936	.913				
-.5	.873	1.000	.764	.768	1.000	1.000	.999	1.000
-.4	.652	.845	.488	.533				
-.3	.391	.369	.217	.291	.953	.998	.888	.937
-.2	.189	.132	.060	.131				
-.1	.081	.060	.010	.049	.176	.134	.051	.157
0	.050	.050	.002	.038	.050	.050	.002	.049
.1	.081	.060	.010	.053	.176	.134	.051	.143
.2	.189	.132	.060	.151				
.3	.391	.369	.217	.318	.953	.998	.888	.948
.4	.652	.845	.488	.524				
.5	.873	1.000	.764	.750	1.000	1.000	.999	1.000
.6	.978	1.000	.936	.917				
.7	.999	1.000	.993	.982	1.000	1.000	1.000	1.000
.8	1.000	1.000	.999	.999				
.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE A-32

POWER APPROXIMATIONS: Max-χ^2 Test
ERROR MODEL 1: J=3, K=4, α=.05

M(1) = Taylor-series approximation   M(2) = Simulated power

ϕ	M(1)	M(2)	M(1)	M(2)
-.9	1.000	1.000	1.000	1.000
-.8	1.000	.999		
-.7	.999	.984	1.000	1.000
-.6	.987	.942		
-.5	.989	.817	1.000	1.000
-.4	.671	.550		
-.3	.388	.311	.969	.953
-.2	.176	.112		
-.1	.075	.048	.174	.166
0	.050	.040	.050	.052
.1	.075	.046	.174	.147
.2	.176	.142		
.3	.388	.316	.969	.961
.4	.671	.554		
.5	.898	.806	1.000	1.000
.6	.987	.937		
.7	.999	.986	1.000	1.000
.8	1.000	1.000		
.9	1.000	1.000	1.000	1.000

TABLE A-33

POWER APPROXIMATIONS: Box-Pierce Test
ERROR MODEL 1: J=4, K=4, α=.05

BP(1) = Imhoff numerical integration
BP(2) = Pearson 3-moment extension
BP(3) = Rao square root transformation
BP(4) = Simulated power

	n=50				n=200			
ϕ	BP(1)	BP(2)	BP(3)	BP(4)	BP(1)	BP(2)	BP(3)	BP(4)
-.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-.8	1.000	1.000	.999	.996				
-.7	.999	1.000	.995	.984	1.000	1.000	1.000	1.000
-.6	.973	1.000	.926	.891				
-.5	.848	.999	.700	.713	1.000	1.000	1.000	1.000
-.4	.607	.698	.391	.478				
-.3	.345	.312	.157	.236	.947	.995	.888	.926
-.2	.159	.114	.044	.107				
-.1	.069	.055	.008	.044	.161	.123	.047	.118
0	.050	.050	.002	.038	.050	.050	.002	.056
.1	.092	.068	.008	.074	.190	.147	.047	.164
.2	.218	.164	.044	.167				
.3	.441	.446	.057	.320	.960	.998	.888	.938
.4	.706	.878	.391	.522				
.5	.906	1.000	.700	.776	1.000	1.000	1.000	1.000
.6	.987	1.000	.926	.921				
.7	1.000	1.000	.995	.986	1.000	1.000	1.000	1.000
.8	1.000	1.000	.999	.998				
.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE A-34

POWER APPROXIMATIONS: Max-χ^2 Test
ERROR MODEL 1: J=4, K=4, α=.05

M(1) = Taylor-series approximation

M(2) = Simulated power

	n=50		n=200	
Ø	M(1)	M(2)	M(1)	M(2)
-.9	1.000	1.000	1.000	1.000
-.8	1.000	.999		
-.7	.999	.993	1.000	1.000
-.6	.987	.937		
-.5	.895	.764	1.000	1.000
-.4	.664	.538		
-.3	.376	.269	.968	.949
-.2	.165	.110		
-.1	.068	.043	.168	.140
0	.050	.040	.050	.059
.1	.085	.063	.183	.156
.2	.198	.136		
.3	.418	.300	.970	.958
.4	.698	.529		
.5	.911	.792	1.000	1.000
.6	.990	.937		
.7	.999	.991	1.000	1.000
.8	1.000	.997		
.9	1.000	1.000	1.000	1.000

TABLE A-35

TAYLOR-SERIES APPROXIMATION
TERM BY TERM: T_0-T_6

n=50 (J=2, K=4) α=.05

Ø	T_0	T_1	T_2	T_3	T_4	T_5	T_6	Sum(M_o)
-.9	.000000	.000000	.000000	.000000	.000000	.000000	.000000	.000000
-.7	.000055	.000258	.000343	-.000092	-.000451	-.000093	.000178	.000198
-.5	.082481	.046790	-.013580	-.001249	-.002504	-.000290	-.000368	.111281
-.3	.611447	.016308	.003511	-.001047	-.000088	-.000186	-.000057	.629888
-.1	.931253	.000099	.000349	.000004	-.000005	.000000	.000000	.931710
.1	.913725	.000097	.000542	.000004	-.000008	.000000	.000000	.914375
.3	.563102	.015018	.008289	-.000919	-.000012	-.000172	.000090	.585396
.5	.064378	.036520	-.006319	-.000165	-.002013	-.000226	-.000249	.091926
.7	.000029	.000135	.000186	-.000039	-.000228	-.000048	.000091	.000125
.9	.000000	.000000	.000000	.000000	.000000	.000000	.000000	.000000

TABLE A-36

TAYLOR-SERIES APPROXIMATION
TERM BY TERM: T_0-T_6
n=200 (J=2, K=4) α=.05

Ø	T_0	T_1	T_2	T_3	T_4	T_5	T_6	Sum(M_o)
-.9	.000000	.000000	.000000	.000000	.000000	.000000	.000000	.000000
-.7	.000000	.000000	.000000	.000000	.000000	.000000	.000000	.000000
-.5	.000000	.000000	.000000	.000000	.000000	.000000	.000000	.000000
-.3	.028312	.008087	-.004553	.000217	.000064	.000108	.000032	.032368
-.1	.831606	.000569	.000513	.000002	-.000005	.000000	.000000	.832685
.1	.815953	.000558	.000681	.000002	-.000002	.000000	.000000	.817193
.3	.026166	.007447	-.003958	.000222	.000056	.000100	.000360	.030069
.5	.000000	.000000	.000000	.000000	.000000	.000000	.000000	.000000
.7	.000000	.000000	.000000	.000000	.000000	.000000	.000000	.000000
.9	.000000	.000000	.000000	.000000	.000000	.000000	.000000	.000000

TABLE A-37

TAYLOR-SERIES APPROXIMATION
TERM BY TERM: T_0-T_6

n=50 (J=3, K=4) α=.05

Ø	T_0	T_2	T_3	T_4	T_5	T_6	Sum(M_o)
-.9	.000000	.000000	.000000	.000000	.000000	.000000	.000000
-.7	.000180	.000049	-.000002	.000004	.000000	.000010	.000241
-.5	.096589	.004883	-.000092	.000085	-.000005	.000306	.101766
-.3	.607383	.003912	-.000020	.000418	-.000001	.000060	.611753
-.1	.924005	.000332	.000000	.000010	.000000	.000000	.924348
.1	.924005	.000332	.000000	.000010	.000000	.000000	.924348
.3	.607383	.003912	-.000020	.000418	-.000001	.000060	.611753
.5	.096589	.004883	-.000092	.000085	.000005	.000306	.101766
.7	.000180	.000049	-.000002	.000004	.000000	.000010	.000241
.9	.000000	.000000	.000000	.000000	.000000	.000000	.000000

NOTE: T_1 = .000000 for all values of Ø.

TABLE A-38

TAYLOR-SERIES APPROXIMATION
TERM BY TERM: T_0-T_6
n=200 (J=3, K=4) α=.05

Ø	T_0	T_2	T_3	T_4	T_6	Sum(M_o)
-.9	.000000	.000000	.000000	.000000	.000000	.000000
-.7	.000000	.000000	.000000	.000000	.000000	.000000
-.5	.000000	.000000	.000000	.000000	.000000	.000000
-.3	.030968	.000199	.000001	.000021	.000003	.031190
-.1	.825715	.000297	.000000	.000009	.000000	.826021
.1	.825715	.000297	.000001	.000009	.000000	.826021
.3	.030968	.000199	.000000	.000021	.000003	.031190
.5	.000000	.000000	.000000	.000000	.000000	.000000
.7	.000000	.000000	.000000	.000000	.000000	.000000
.9	.000000	.000000	.000000	.000000	.000000	.000000

NOTE: $T_1 = T_5 = .000000$ for all values of Ø.

TABLE A-39

TAYLOR-SERIES APPROXIMATION
TERM BY TERM: T_0-T_8
n=50 (J=4, K=4) α=.05

Ø	T_0	T_2	T_4	T_6	T_8	Sum(M_o)
-.9	.000000	.000000	.000000	.000000	.000000	.000000
-.7	.000211	.000021	.000001	.000000	.000005	.000238
-.5	.102061	.002087	.000058	.000166	.000409	.104782
-.3	.621144	.001826	.000219	.000046	.000004	.623240
-.1	.931275	.000166	.000005	.000000	.000000	.931446
.1	.914089	.000163	.000005	.000000	.000000	.914257
.3	.579597	.001704	.000205	.000043	.000003	.581553
.5	.086720	.001773	.000049	.000141	.000348	.089032
.7	.000148	.000014	.000001	.000000	.000004	.000168
.9	.000000	.000000	.000000	.000000	.000000	.000000

NOTE: $T_1 = T_3 = T_5 = T_7 = .000000$ for all values of Ø.

TABLE A-40

TAYLOR-SERIES APPROXIMATION
TERM BY TERM: T_0-T_8

n=200 (J=4, K=4) α=.05

Ø	T_0	T_2	T_4	T_6	Sum(M_o)
-.9	.000000	.000000	.000000	.000000	.000000
-.7	.000000	.000000	.000000	.000000	.000000
-.5	.000000	.000000	.000000	.000000	.000000
-.3	.031669	.000093	.000011	.000002	.031776
-.1	.832211	.000148	.000005	.000000	.832364
.1	.816853	.000145	.000005	.000002	.817003
.3	.029551	.000087	.000010	.000000	.029651
.5	.000000	.000000	.000000	.000000	.000000
.7	.000000	.000000	.000000	.000000	.000000
.9	.000000	.000000	.000000	.000000	.000000

NOTE: $T_1 = T_3 = T_5 = T_7 = T_8 = .000000$ for all values of Ø.

TABLE A-41

ASYMPTOTIC POWERS

K=4, J=1, α=.05

Lag Parameter:	-.5		-.3		-.1		.1		.3		.5	
Sample Size	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX
50	.865	.929	.398	.407	.082	.076	.082	.076	.398	.407	.865	.929
100	.996	.999	.687	.738	.112	.105	.112	.105	.687	.738	.996	.999
200	1.000	1.000	.951	.973	.177	.175	.177	.175	.951	.973	1.000	1.000
500	1.000	1.000	1.000	1.000	.401	.424	.401	.424	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	.718	.761	.718	.761	1.000	1.000	1.000	1.000
2000	1.000	1.000	1.000	1.000	.964	.978	.964	.978	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

K=4, J=2, α=.05

Lag Parameter:	-.5		-.3		-.1		.1		.3		.5	
Sample Size	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX
50	.838	.918	.352	.389	.070	.069	.094	.086	.451	.437	.900	.938
100	.994	.999	.656	.729	.099	.098	.124	.115	.723	.750	.997	.999
200	1.000	1.000	.944	.972	.163	.168	.192	.184	.958	.974	1.000	1.000
500	1.000	1.000	1.000	1.000	.386	.420	.416	.431	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	.708	.759	.728	.764	1.000	1.000	1.000	1.000
2000	1.000	1.000	1.000	1.000	.963	.978	.966	.978	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE A-42

ASYMPTOTIC POWERS

K=4, J=3, α=.05

Lag Parameter:	-.5		-.3		-.1		.1		.3		.5	
Sample Size	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX
50	.873	.903	.391	.393	.081	.076	.081	.076	.391	.393	.873	.903
100	.996	.998	.689	.723	.110	.105	.110	.105	.689	.723	.996	.998
200	1.000	1.000	.953	.969	.176	.174	.176	.174	.953	.969	1.000	1.000
500	1.000	1.000	1.000	1.000	.400	.423	.400	.423	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	.719	.760	.719	.760	1.000	1.000	1.000	1.000
2000	1.000	1.000	1.000	1.000	.965	.978	.965	.978	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

K=4, J=4, α=.05

Lag Parameter:	-.5		-.3		-.1		.1		.3		.5	
Sample Size	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX
50	.848	.898	.345	.379	.069	.069	.092	.086	.441	.421	.906	.913
100	.995	.998	.658	.717	.097	.098	.122	.115	.724	.736	.997	.999
200	1.000	1.000	.947	.968	.161	.168	.190	.183	.960	.970	1.000	1.000
500	1.000	1.000	1.000	1.000	.385	.419	.415	.429	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	.709	.758	.729	.763	1.000	1.000	1.000	1.000
2000	1.000	1.000	1.000	1.000	.963	.977	.966	.978	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE A-43

ASYMPTOTIC POWERS

K=8, J=1, α=.05

Lag Parameter:	-.5		-.3		-.1		.1		.3		.5	
Sample Size	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX
50	.827	.892	.355	.336	.077	.068	.077	.068	.355	.336	.827	.892
100	.990	.998	.600	.658	.097	.087	.097	.087	.600	.658	.990	.998
200	1.000	1.000	.898	.953	.142	.137	.142	.137	.898	.953	1.000	1.000
500	1.000	1.000	1.000	1.000	.307	.345	.307	.345	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	.595	.686	.595	.686	1.000	1.000	1.000	1.000
2000	1.000	1.000	1.000	1.000	.917	.962	.917	.962	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

K=8, J=2, α=.05

Lag Parameter:	-.5		-.3		-.1		.1		.3		.5	
Sample Size	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX
50	.795	.883	.316	.325	.069	.064	.085	.074	.395	.362	.853	.906
100	.987	.998	.565	.652	.088	.083	.105	.093	.631	.671	.990	.998
200	1.000	1.000	.887	.952	.132	.133	.152	.143	.911	.955	1.000	1.000
500	1.000	1.000	1.000	1.000	.295	.341	.319	.349	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	.584	.684	.605	.688	1.000	1.000	1.000	1.000
2000	1.000	1.000	1.000	1.000	.914	.962	.920	.963	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE A-44

ASYMPTOTIC POWERS

K=8, J=3, α=.05

Lag Parameter:	-.5		-.3		-.1		.1		.2		.3	
Sample Size	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX	BP	MX
50	.823	.887	.357	.335	.076	.068	.076	.068	.357	.335	.823	.887
100	.986	.998	.602	.657	.096	.087	.096	.087	.602	.657	.986	.998
200	1.000	1.000	.901	.953	.142	.137	.142	.137	.901	.953	1.000	1.000
500	1.000	1.000	1.000	1.000	.307	.345	.307	.345	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	.595	.686	.595	.686	1.000	1.000	1.000	1.000
2000	1.000	1.000	1.000	1.000	.917	.962	.917	.962	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

BIBLIOGRAPHY

1. Abrahamse, A.P.J. and Koerts, J. (1969). A comparison between the power of the Durbin-Watson test and the power of the BLUS test. Journal of the Ammerican Statistical Association, 64, 938-948.

2. Abrahamse, A.P.J. and Louter, A.S. (1971). On a new test for autocorrelation in least squares regression. Biometrika, 58, 53-60.

3. Abramowitz, M. and Stegun, I.A. (1970). Handbook of Mathematical Functions. U.S. Department of Commerce, Washington, D.C.

4. Anderson, T.W. (1971). The Statistical Analysis of Time Series. Wiley and Sons, New York.

5. Anderson, T.W. (1948). On the theory of testing serial correlation. Skand. Aktuar Tisdkr., 31, 88-116.

6. Bartlett, M.S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. Journal of the Royal Statistical Society, Series B, 8, 27-41.

7. Box, G.E.P. and Pierce, D.A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association, 65, 1509-1526.

8. Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis Forecasting and Control. Holden-Day, San Francisco.

9. Durbin, J. (1969). Tests for serial correlation in regression analysis based on the periodogram of least-squares residuals. Biometrika, 56, 1-15.

10. Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression. I. Biometrika, 27, 409-428.

11. Durbin, J. and Watson, G.S. (1951). Testing for serial correlation in least squares regression. II. Biometrika, 38, 159-178.

12. Durbin, J. and Watson, G.S. (1971). Testing for serial correlation in least squares regression. III. Biometrika, 58, 1-19.

13. Fuller, W.A. (1976). Introduction to Statistical Time Series. Wiley and Sons, New York.

14. Geary, R.C. (1970). Relative efficiency of count of sign changes for assessing residual autoregression in least squares regressions. Biometrika, 57, 123-127.

15. Grenander, U. and Rosenblatt, M. (1957). Statistical Analysis of Stationary Time Series. Wiley and Sons, New York.

16. Hannan, E.J. (1960). Time Series Analysis. Methuen, London.

17. Hannan, E.J.(1957). Testing for serial correlation in least squares regression. Biometrika, 44, 57-66.

18. Hannan, E.J. and Terrell, R.D. (1968). Testing for serial correlation after least squares regression. Econometrica, 36, 133-150.

19. Hart, B.I. (1942). Tabulation of the probabilities for the ratio of the mean square successive difference to the variance. The Annals of Mathematical Statistics, 13, 207-214.

20. Henshaw, R.C. (1966). Testing single equation least squares regression models for autocorrelated disturbances. Econometrica, 34, 646-660.

21. Hodges, J.L. and Lehmann, E.L. (1956). The efficiency of some nonparametric competitors of the t-test. The Annals of Mathematical Statistics, 27, 324-335.

22. Hodges, J.L. and Lehmann, E.L. (1970). Deficiency. The Annals of Mathematical Statistics, 41, 783-801.

23. Imhoff, J.P. (1961). Computing the distribution of quadratic forms in normal variables. Biometrika, 48, 419-426.

24. McClave, J.T. (1978). Estimating the order of autoregressive models: The max-χ^2 method. Journal of the American Statistical Association, 73, 122-128.

25. Pearson, E.S. (1959). Note on an approximation to the distribution of non-central χ^2. Biometrika, 46, 264.

26. Phillips, G.D.A. and Harvey, A.C. (1974). A simple test for serial correlation in regression analysis. Journal of the American Statistical Association, 69, 935-939.

27. Rao, C.R. (1973). Linear Statistical Inference and its Applications. Wiley and Sons, New York.

28. Scheffe´, H. (1959). The Analysis of Variance. Wiley and Sons, New York.

29. Schmidt, P. (1972). A generalization of the Durbin-Watson test. Australian Economic Papers, 11, 203-209.

30. Smith, V.K. (1976). The estimated power of several tests for autocorrelation with non-first-order alternative. Journal of the American Statistical Association, 71, 879-883.

31. Tan, W.Y. (1977). On the distribution of quadratic forms in normal random variables. The Canadien Journal of Statistics 5, 241-250.

32. Theil, H. (1965). The analysis of disturbances in regression analysis. Journal of the American Statistical Association, 60, 1067-1079.

33. Theil, H. and Nagar, A.L. (1961). Testing the independence of regression disturbances. Journal of the American Statistical Association, 56, 793-806.

34. Wallis, K.F. (1972). Testing for fourth order autocorrelation in quarterly regression equations. Econometrica, 40, 617-636.

BIOGRAPHICAL SKETCH

Terence L. Sincich was born on July 30, 1953, in Sharon, Pennsylvania. Terry's family moved to Phoenix, Arizona,in 1978 where they presently make their home. Upon graduating from Hickory High School in 1971, Terry matriculated to Grove City College in Grove City, Pennsylvania. There he received a Bachelor of Science degree in mathematics in 1975.

Terry came to the University of Florida in the fall of 1975 and received a Master of Statistics degree in 1977. Since that time he has been pursuing a doctoral degree in statistics. While studying at the University of Florida, Terry has been employed by the Department of Statistics as a teaching and research assistant.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

J. T. McClave, Chairman
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

M. C. K. Yang
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

D. J. Nye
Associate Professor of Finance

This dissertation was submitted to the Graduate Faculty of the Department of Statistics in the College of Liberal Arts and Sciences and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1980

Dean, Graduate School

UNIVERSITY OF FLORIDA
3 1262 08666 923 0